

FIG. 1

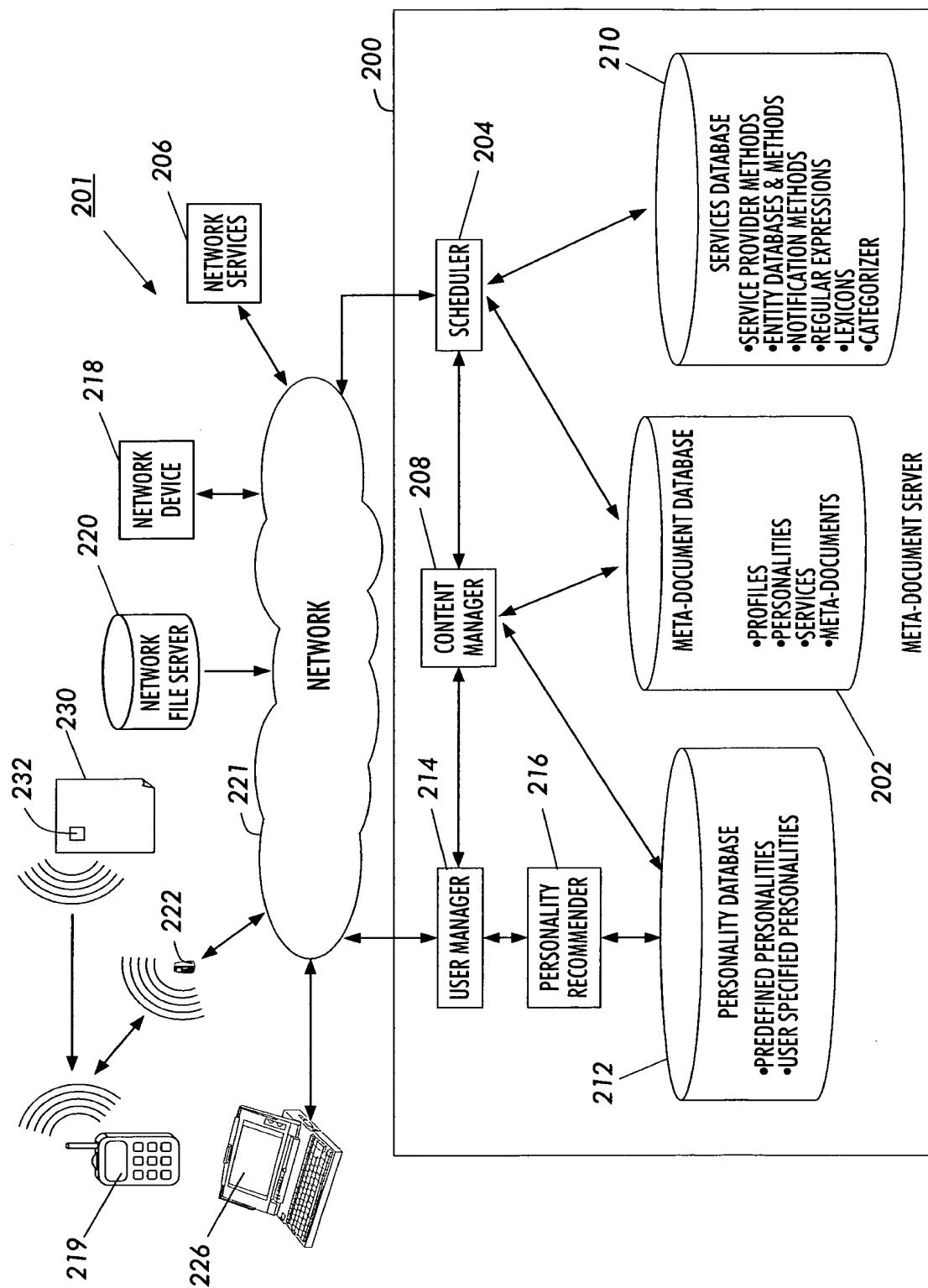


FIG. 2

3/69

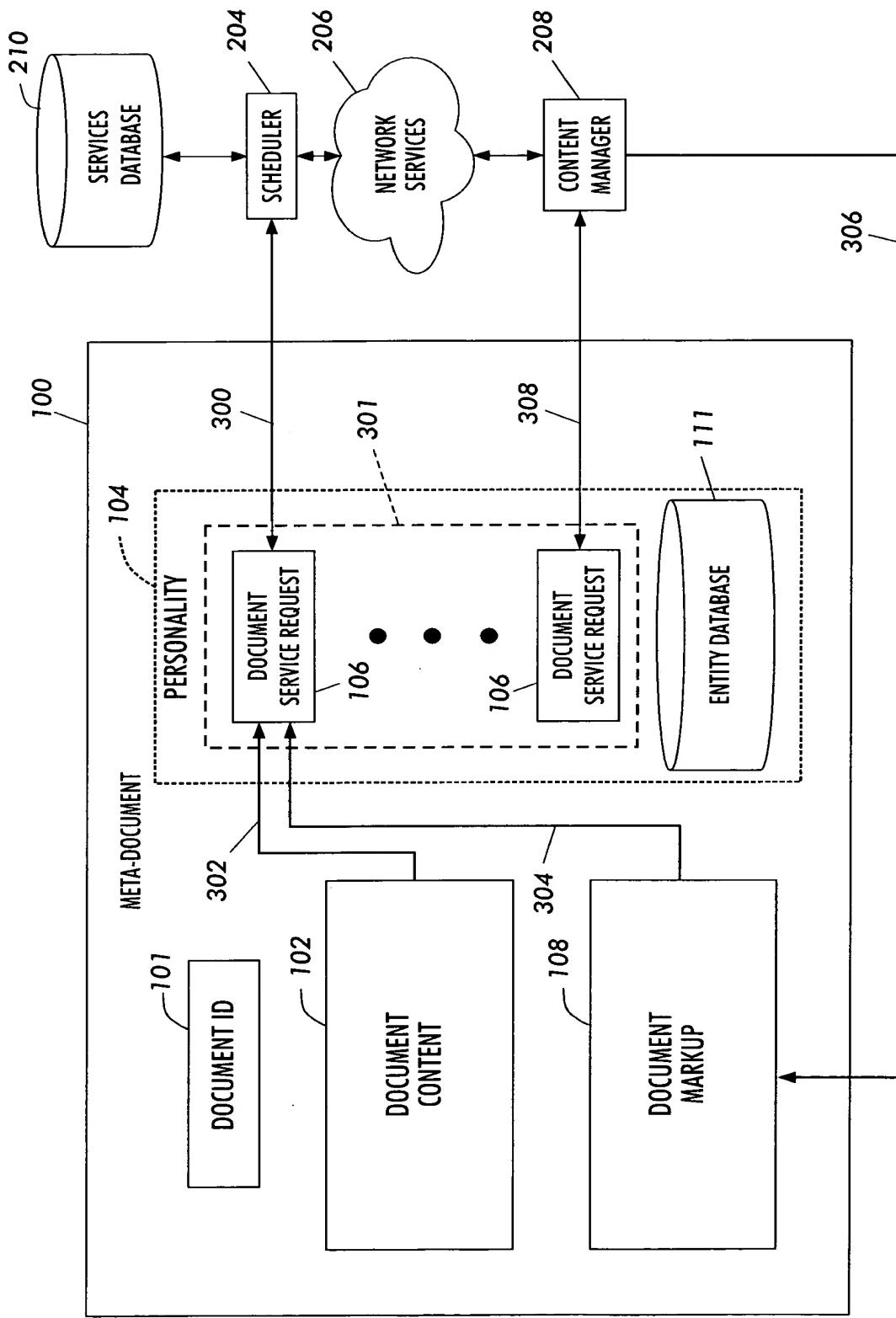
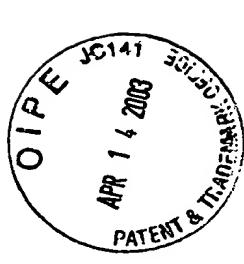
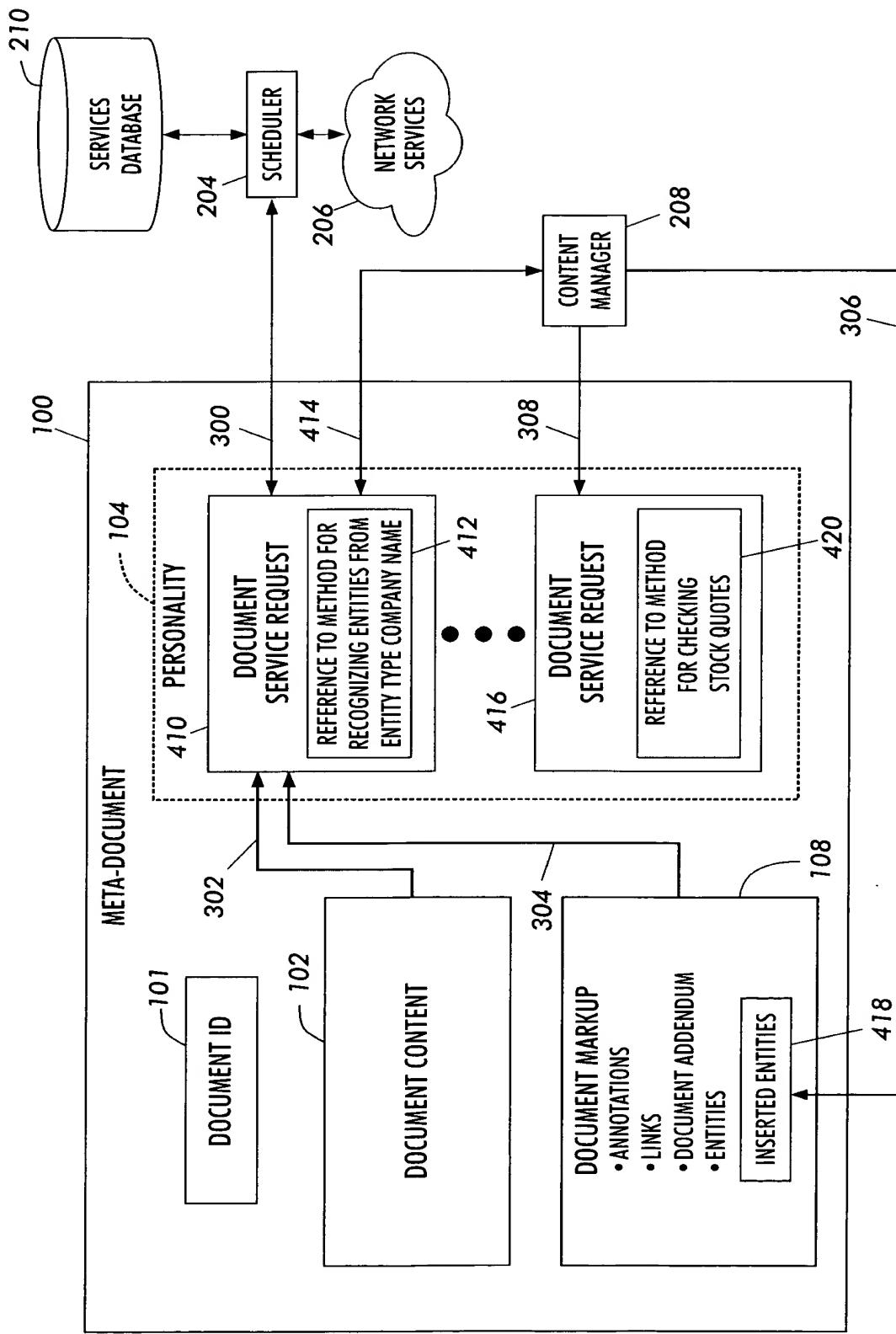


FIG. 3



2176  
09/683,240

4/69



5/69

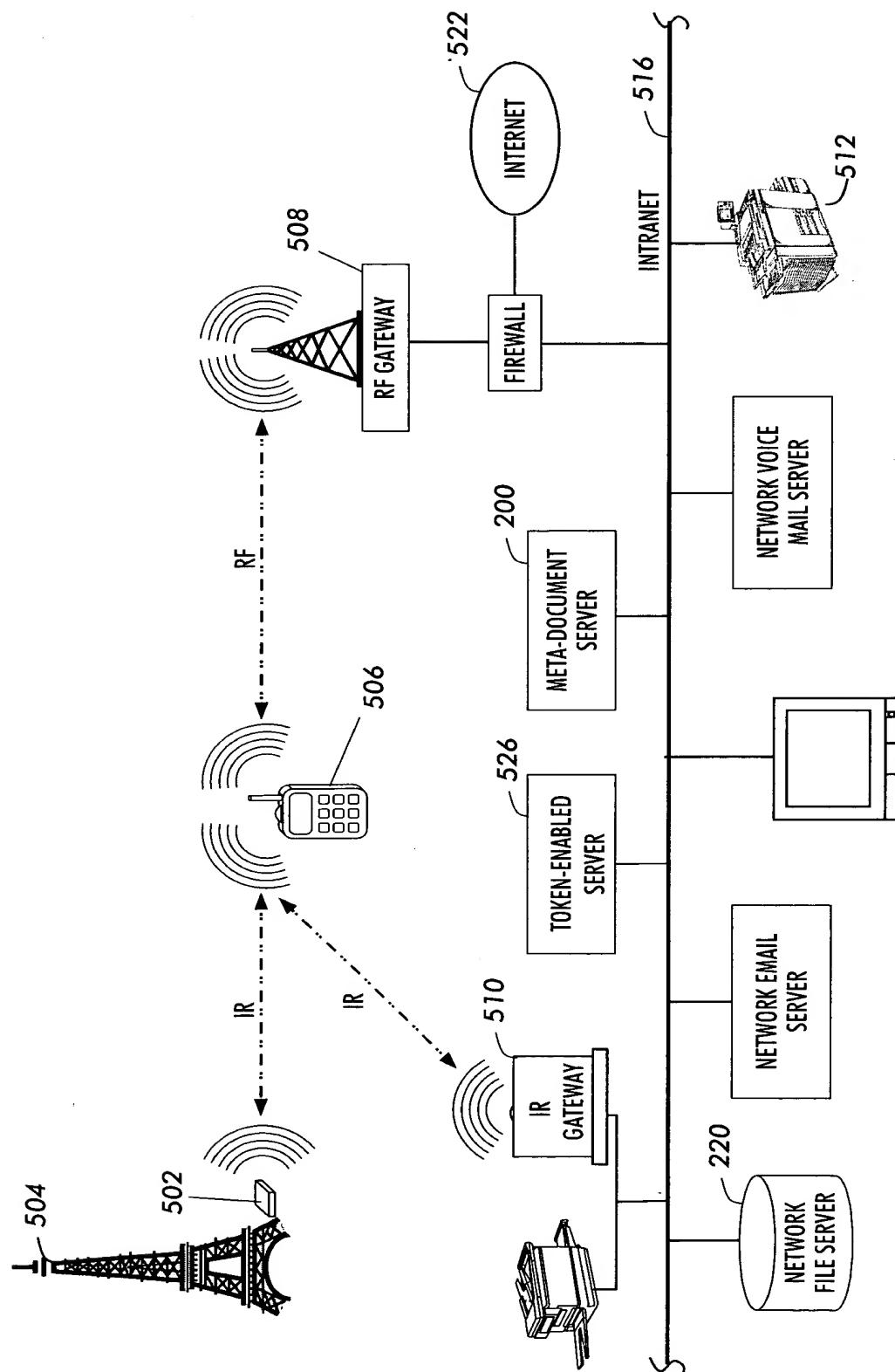


FIG. 5

6/69

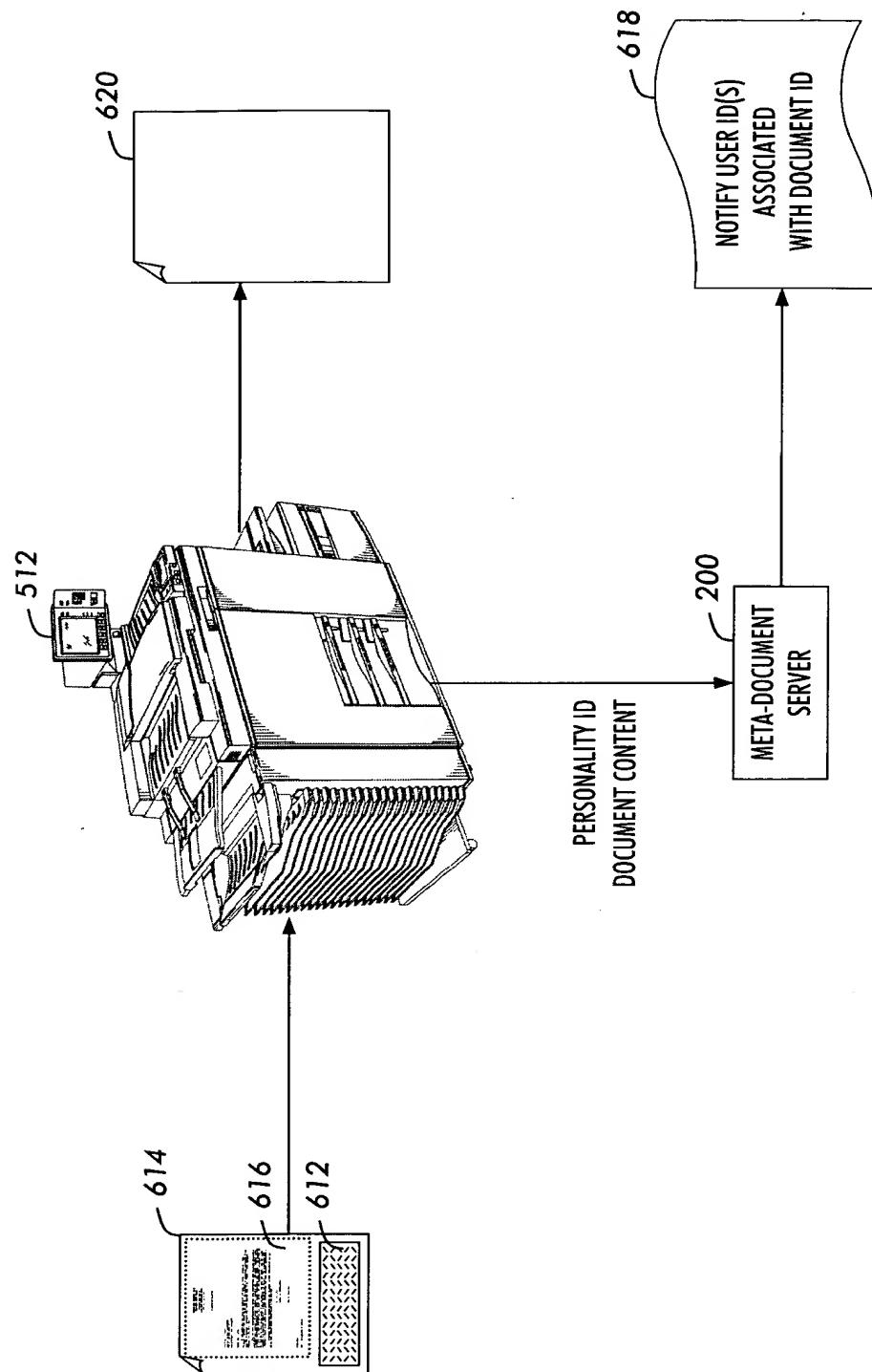


FIG. 6

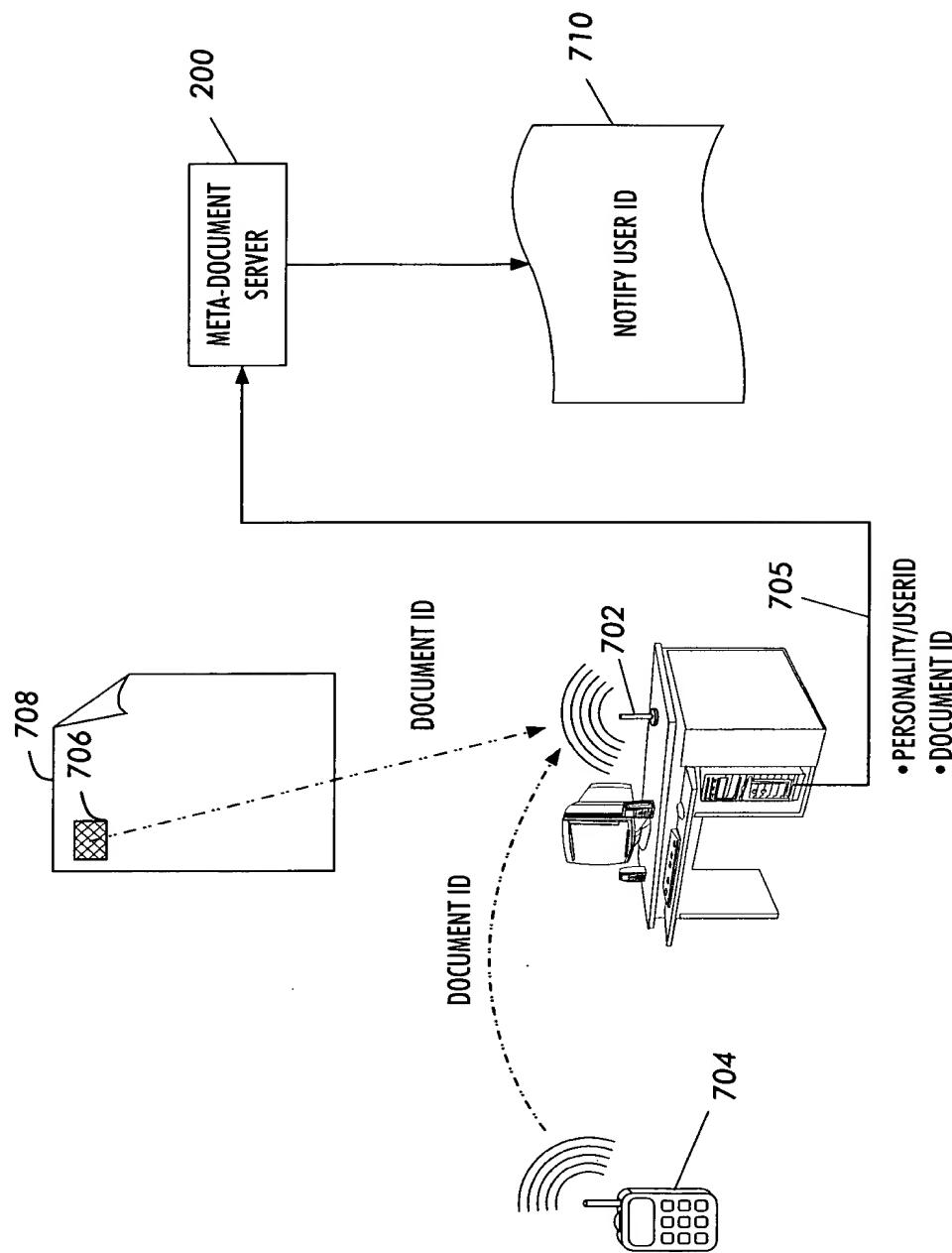
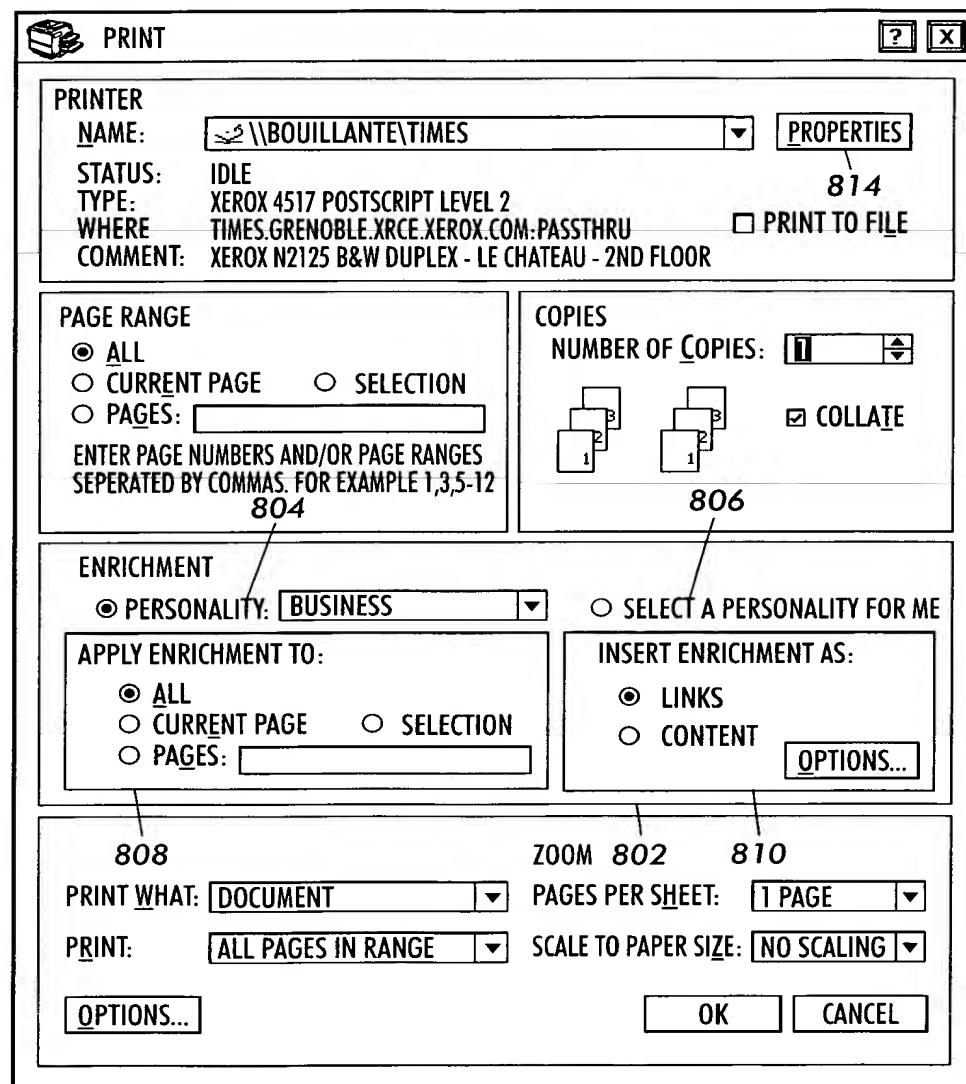


FIG. 7



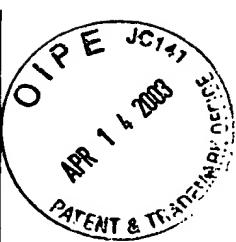
2176  
09/683,240

8/69



**FIG. 8**

800



2176  
09/683,240

9/69

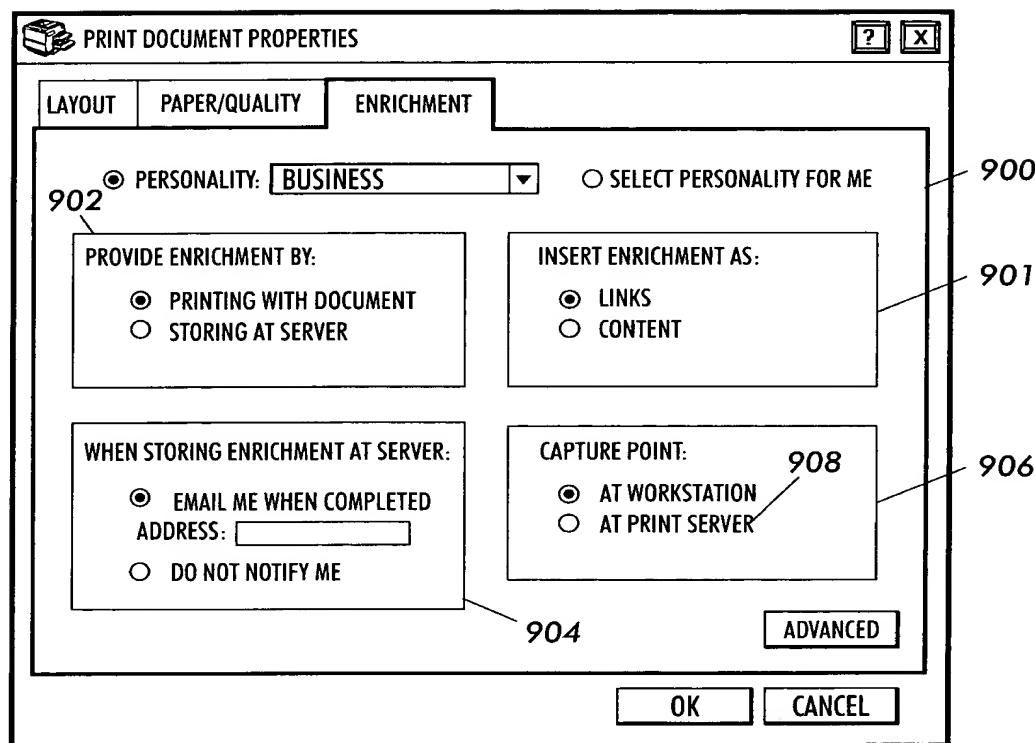
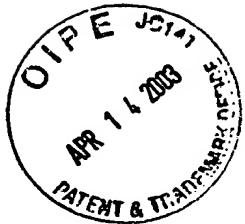


FIG. 9



2176  
09/683,240

10/69

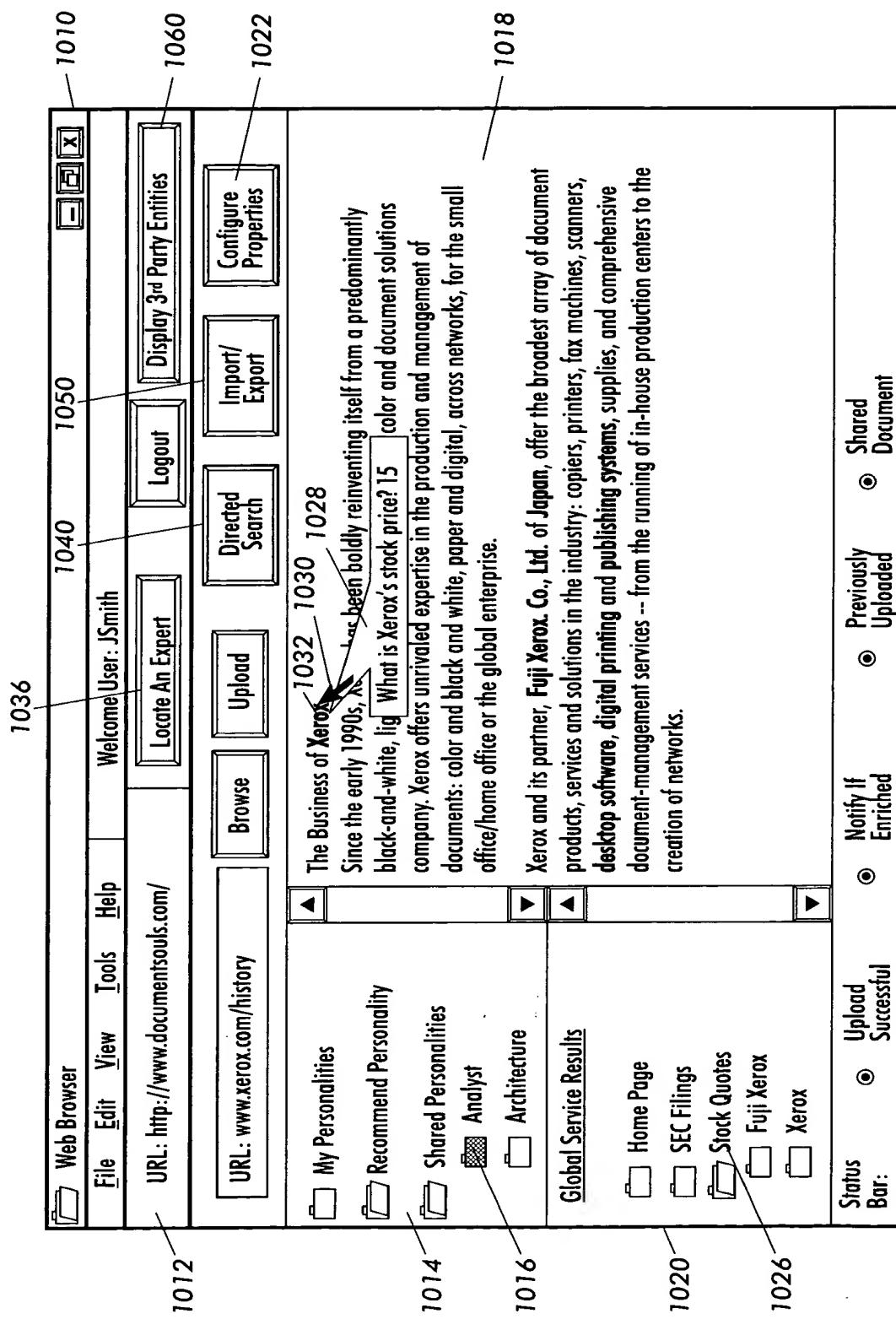
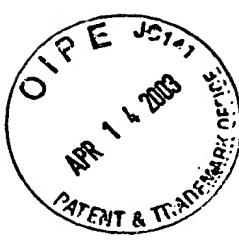


FIG. 10



2176  
09/683,240

11/69

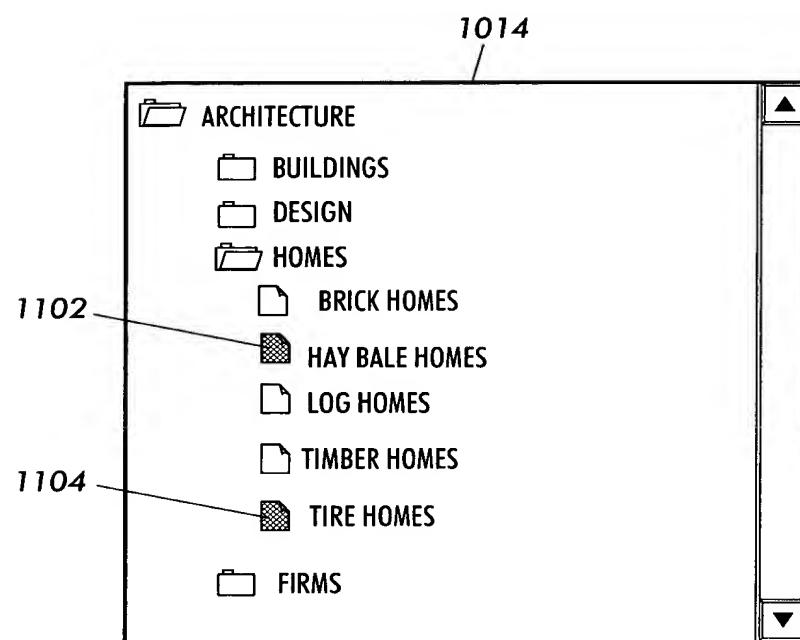


FIG. 11

12/69

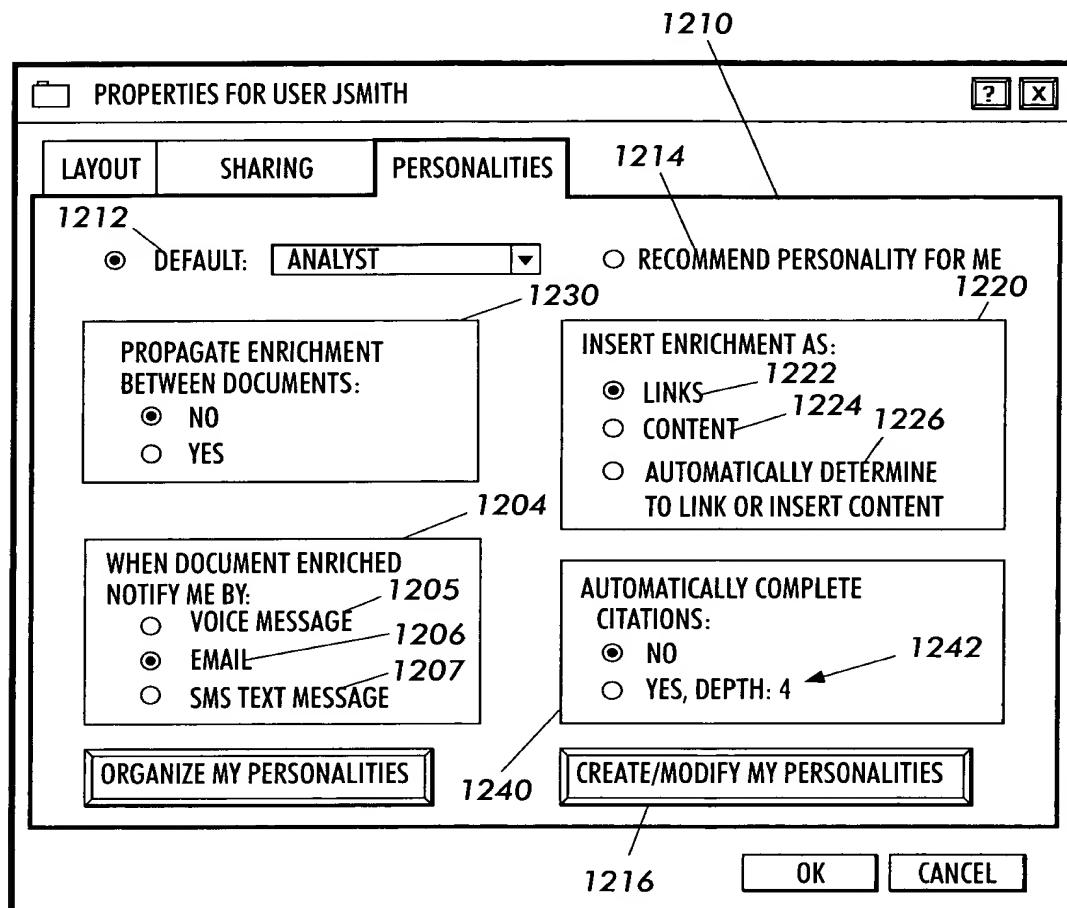
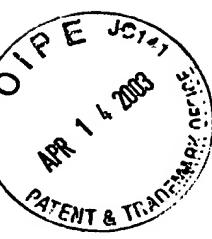


FIG. 12



2176  
09/683,240

13/69

1310

CREATE/MODIFY A GROUP OF EXISTING PERSONALITIES

LIST OF AVAILABLE PERSONALITIES

1316

Name	Creation Date
Technology Watch in Computer Sciences	2001-03-16 14:08:06
Watch Business	2001-03-16 14:08:08
Watch Legal	2001-03-16 14:08:10
Watch Intellectual Property	2001-03-16 14:08:10
Scientific Watch in Medicine	2001-03-16 14:08:10

1312

1318

1314

1319

1318

Click on this icon to view all available service categories

1319

those Services are classified by Service Categories, a ticked checkbox meaning that the according service is attached to the selected personality.

1314

Search Engine

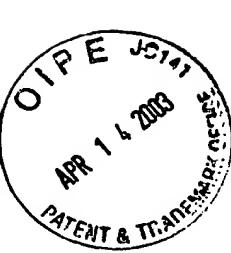
- luckygoogle People Google Lucky Result

Stock Quotes

- bloomberg Stock Quotes

OK CANCEL

FIG. 13



2176  
09/683,240

14/69

SEL	NAME	ENTITY TYPE	ANNOTATE	DEPTH	PROVIDER	COST	QUERY	REFRESH
<input checked="" type="checkbox"/>	COMPUTING	COMPUTING	ALWAYS	0	YAHOO	FREE	LINK	DAILY
<input type="checkbox"/>	BIOLOGY	BIOLOGY	ALWAYS	0	GOOGLELUCKY	NULL	LINK	NEVER
<input type="checkbox"/>	MEDICINE	MEDICINE	ALWAYS	0	ALTAVISTA	FREE	LINK	AUTO
<input type="checkbox"/>	APPLY PAIRS OF ENTITIES AND IDENTIFY 3RD PARTY ENTITIES	COMPANIES	EXPERT	0	GOOGLELUCKY	NULL	CONTENT	NULL
<input type="checkbox"/>	APPLY PAIRS OF ENTITIES USING GOOGLELUCKY	PRODUCTS	ALWAYS	2	NORTHERNLIGHTS	NULL	LINK	WEEKLY
<input type="checkbox"/>	COUNTRIES	COUNTRIES	NOVICE	0	GOOGLELUCKY	FREE	CONTENT	UNTIL 2003
<input type="checkbox"/>	ASTRONOMY	ASTRONOMY	ALWAYS	1	GOOGLE	NULL	LINK	MONTHLY
<input checked="" type="checkbox"/>	PEOPLE	PEOPLE	ALWAYS	3	GOOGLE.PEOPLE	NULL	LINK	NULL

1424

1432

VIEW HIERARCHY  
SET DEPTH

FIG. 14

15/69

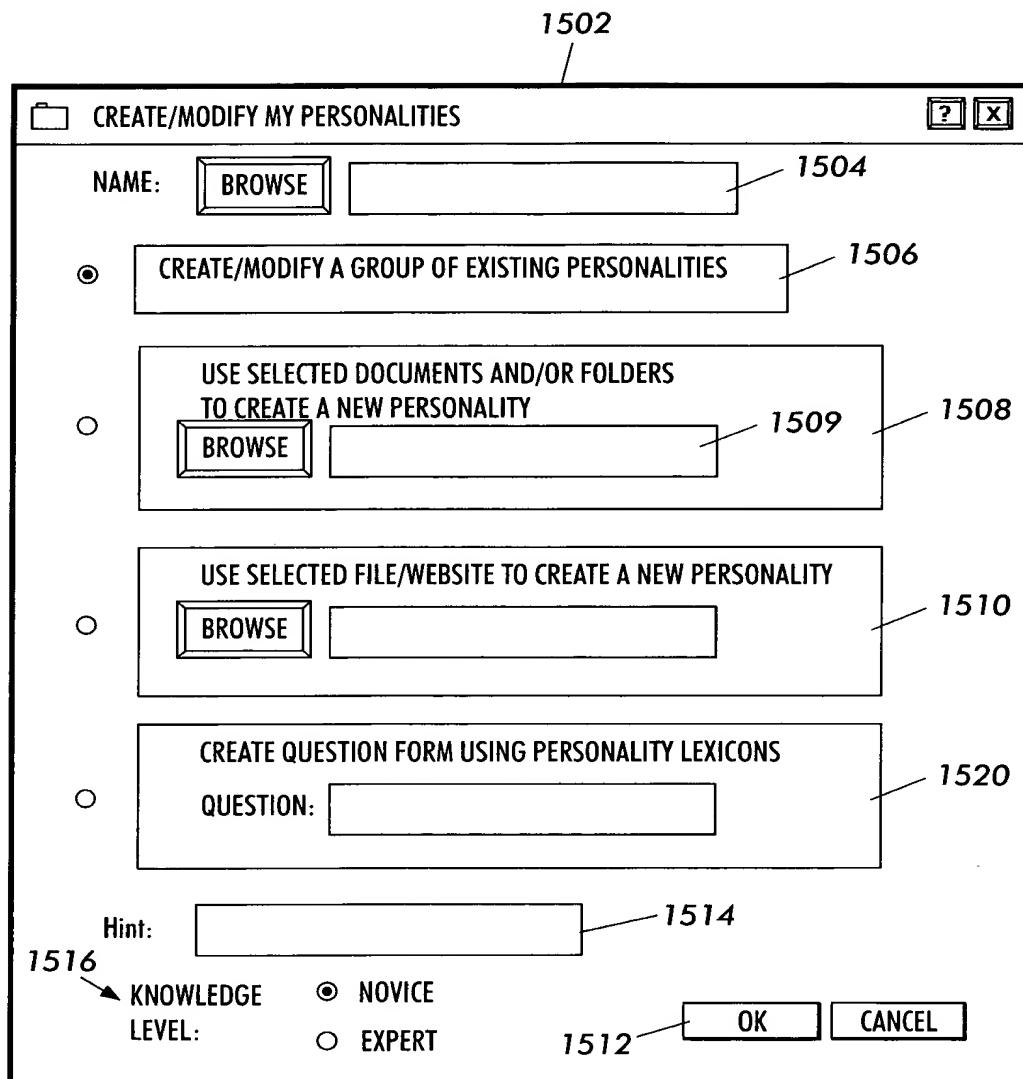


FIG. 15



2176  
09/683,240

16/69

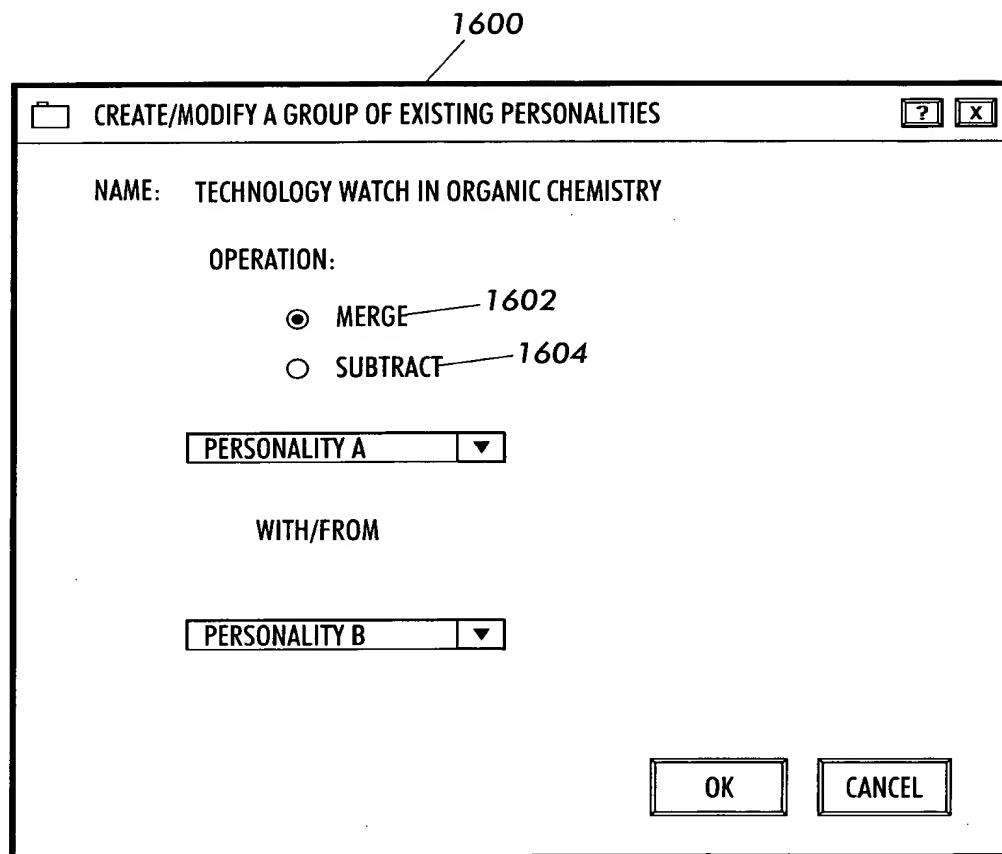


FIG. 16

17/69

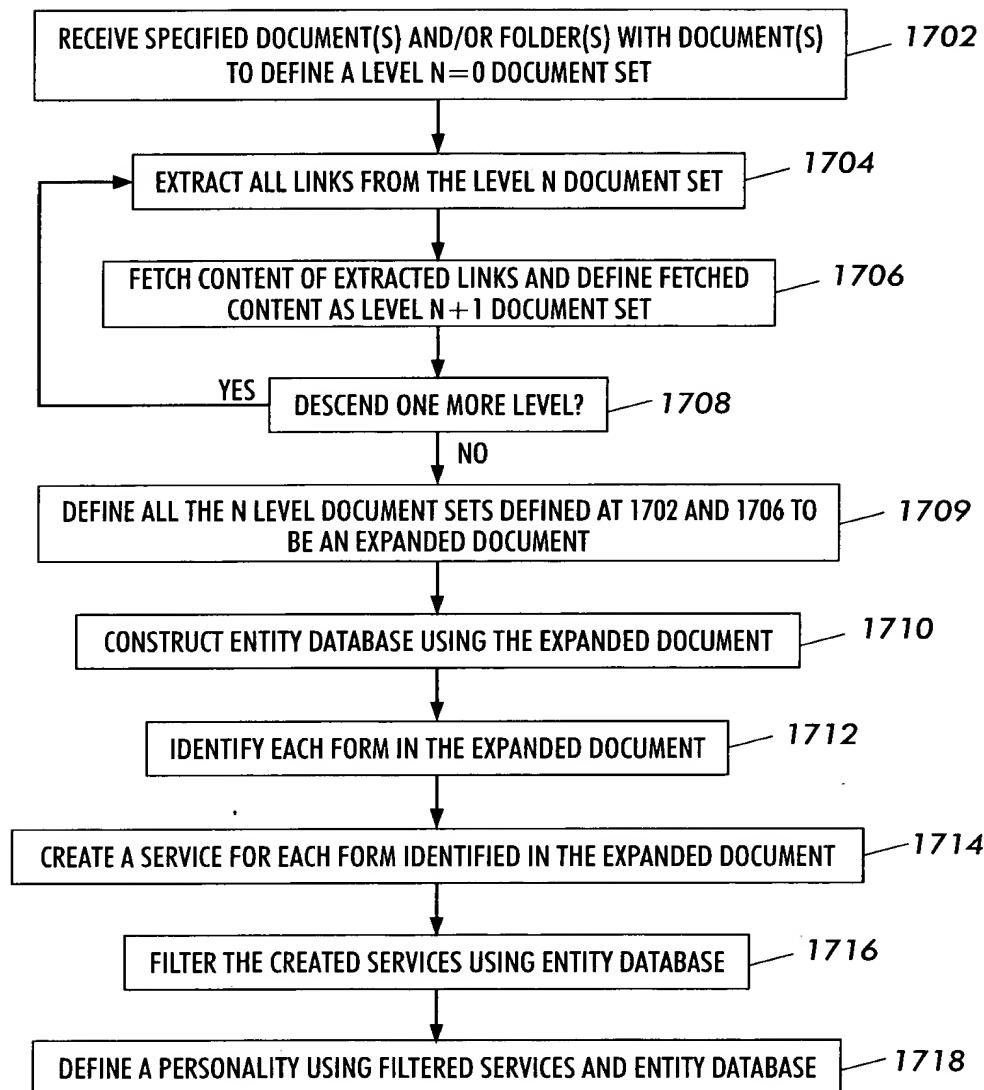


FIG. 17



2176  
09/683,240

18/69

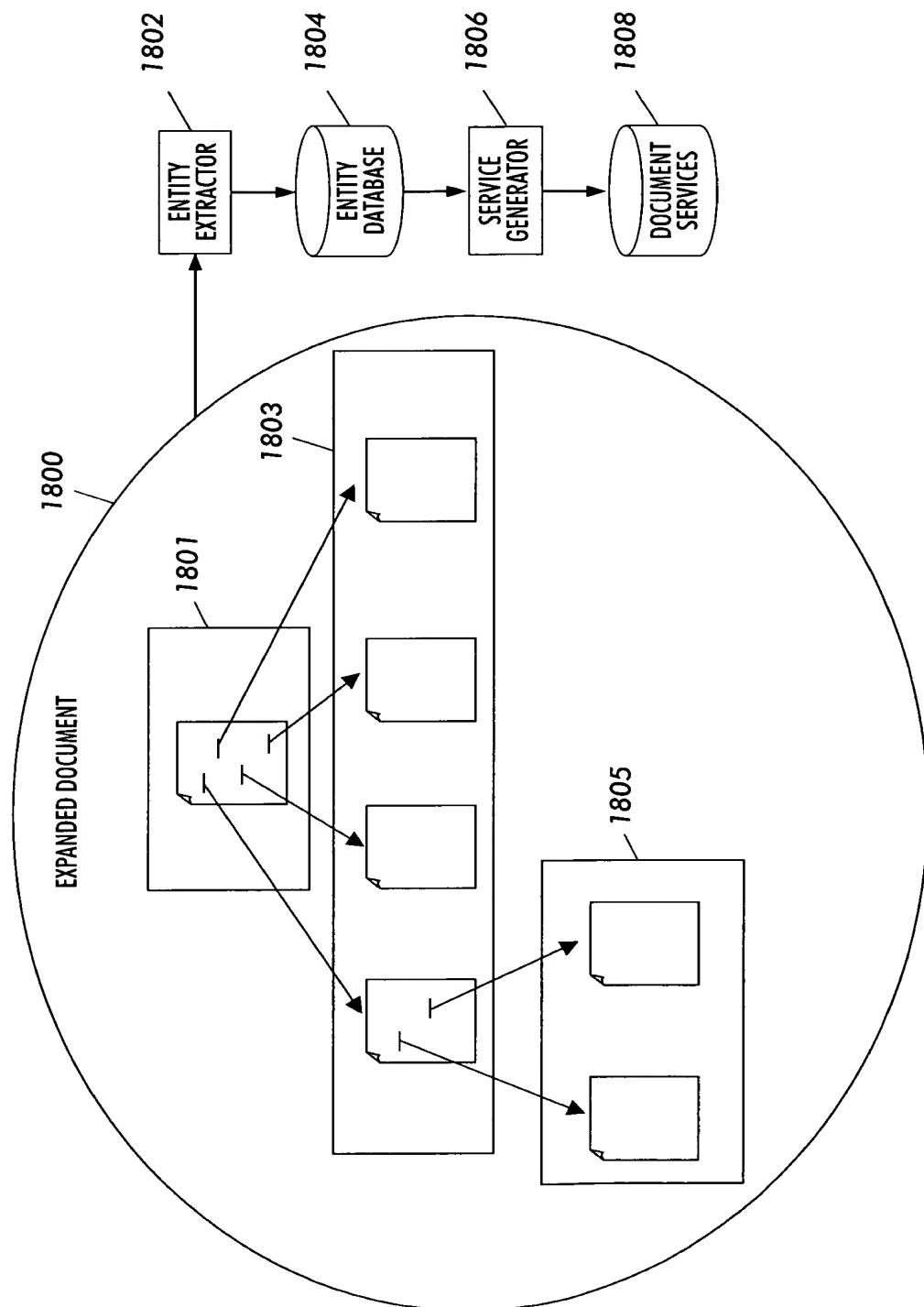


FIG. 18



2176  
09/683,240

19/69

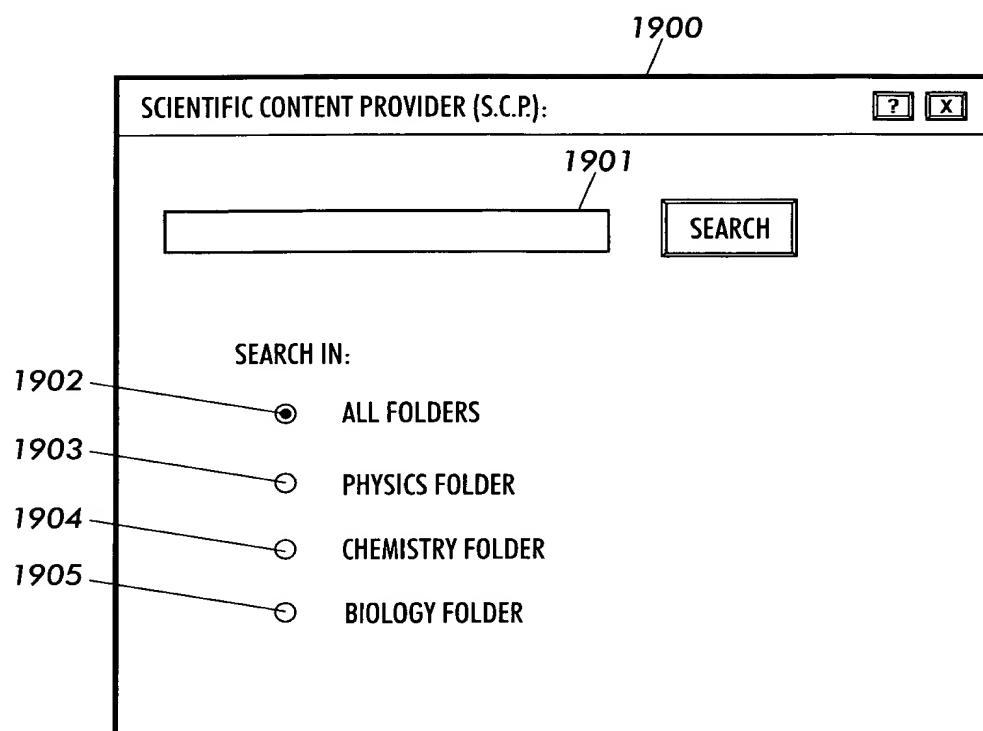
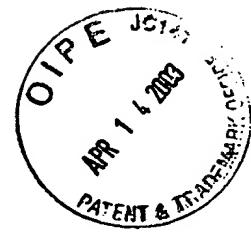


FIG. 19



2176  
09/683,240

20/69

SCIENTIFIC SEARCH □ ▶						
SEL	NAME	ENTITY TYPE	ANNOTATE	DEPTH	PROVIDER	COST
<input checked="" type="checkbox"/>	ALL FOLDERS RESULT OF SCIENTIFIC CONTENT PROVIDER	STANDARD	ALWAYS	0	S. C. P.	FREE
<input type="checkbox"/>	PHYSICS RESULT OF SCIENTIFIC CONTENT PROVIDER	PHYSICS	ALWAYS	0	S. C. P.	FREE
<input type="checkbox"/>	CHEMISTRY RESULT OF SCIENTIFIC CONTENT PROVIDER	CHEMISTRY	ALWAYS	0	S. C. P.	FREE
<input type="checkbox"/>	BIOLOGY RESULT OF SCIENTIFIC CONTENT PROVIDER	BIOLOGY	ALWAYS	0	S. C. P.	FREE

2002  
2004

FIG. 20

21/69

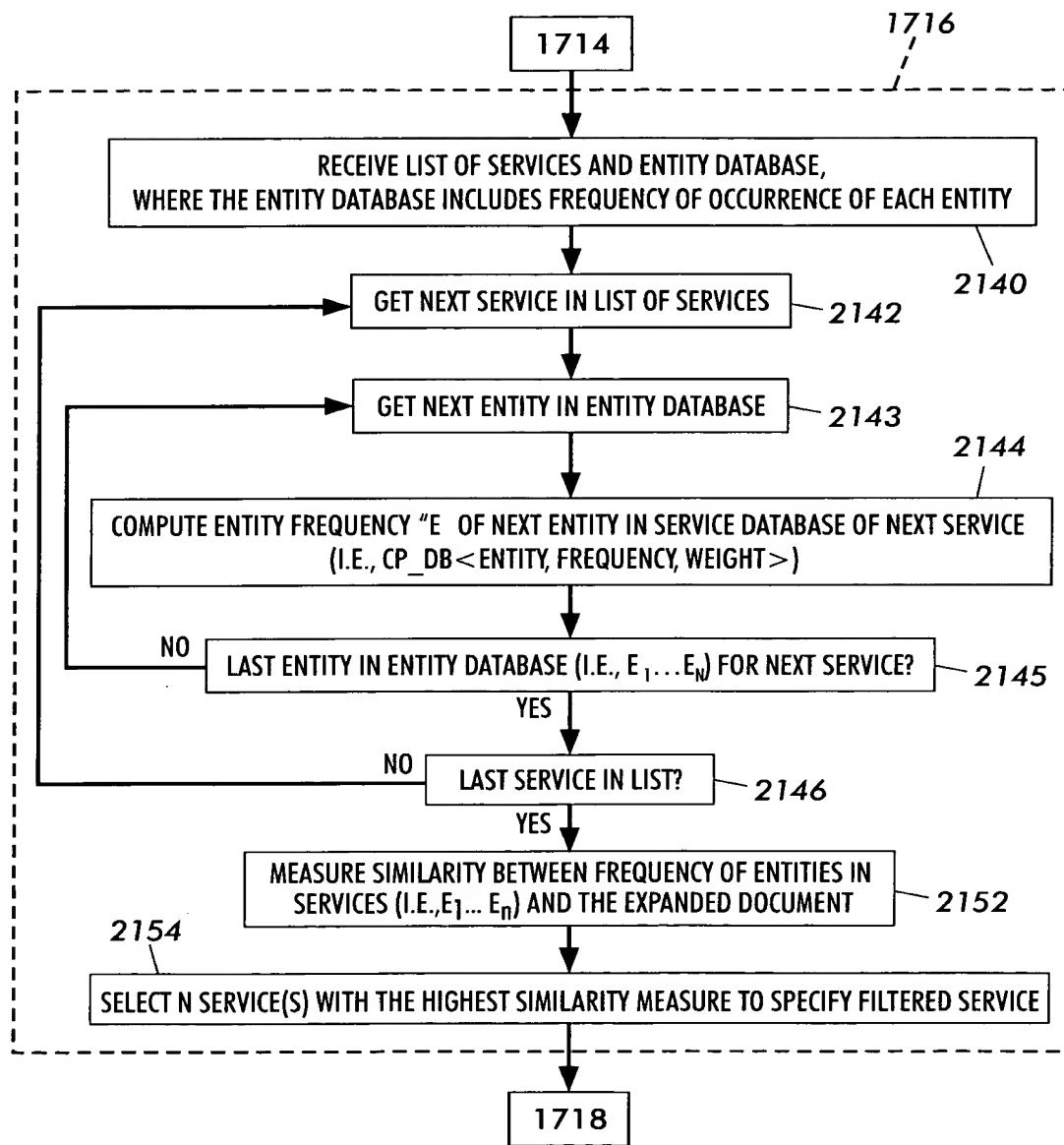
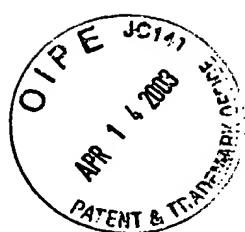
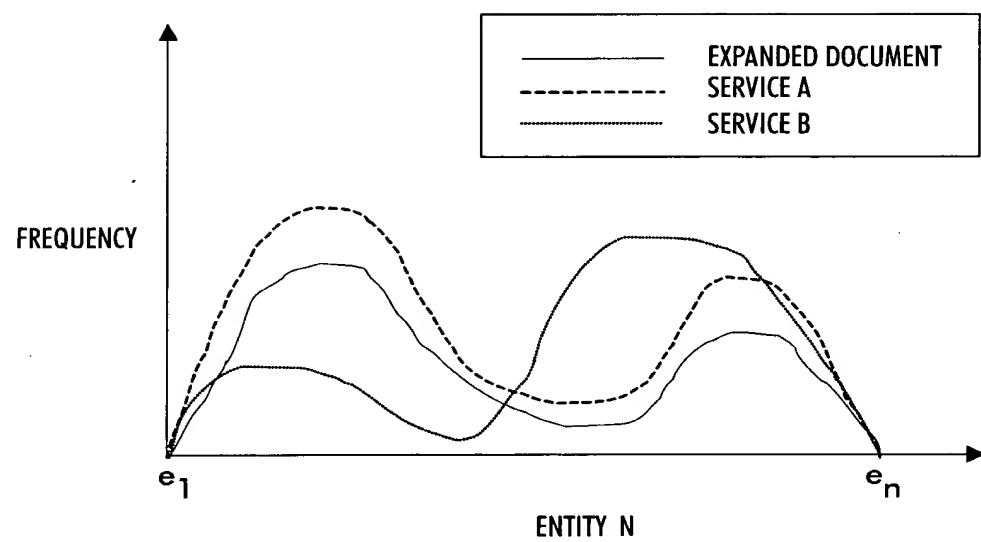


FIG. 21



2176  
09/683,240

22/69



**FIG. 22**

23/69

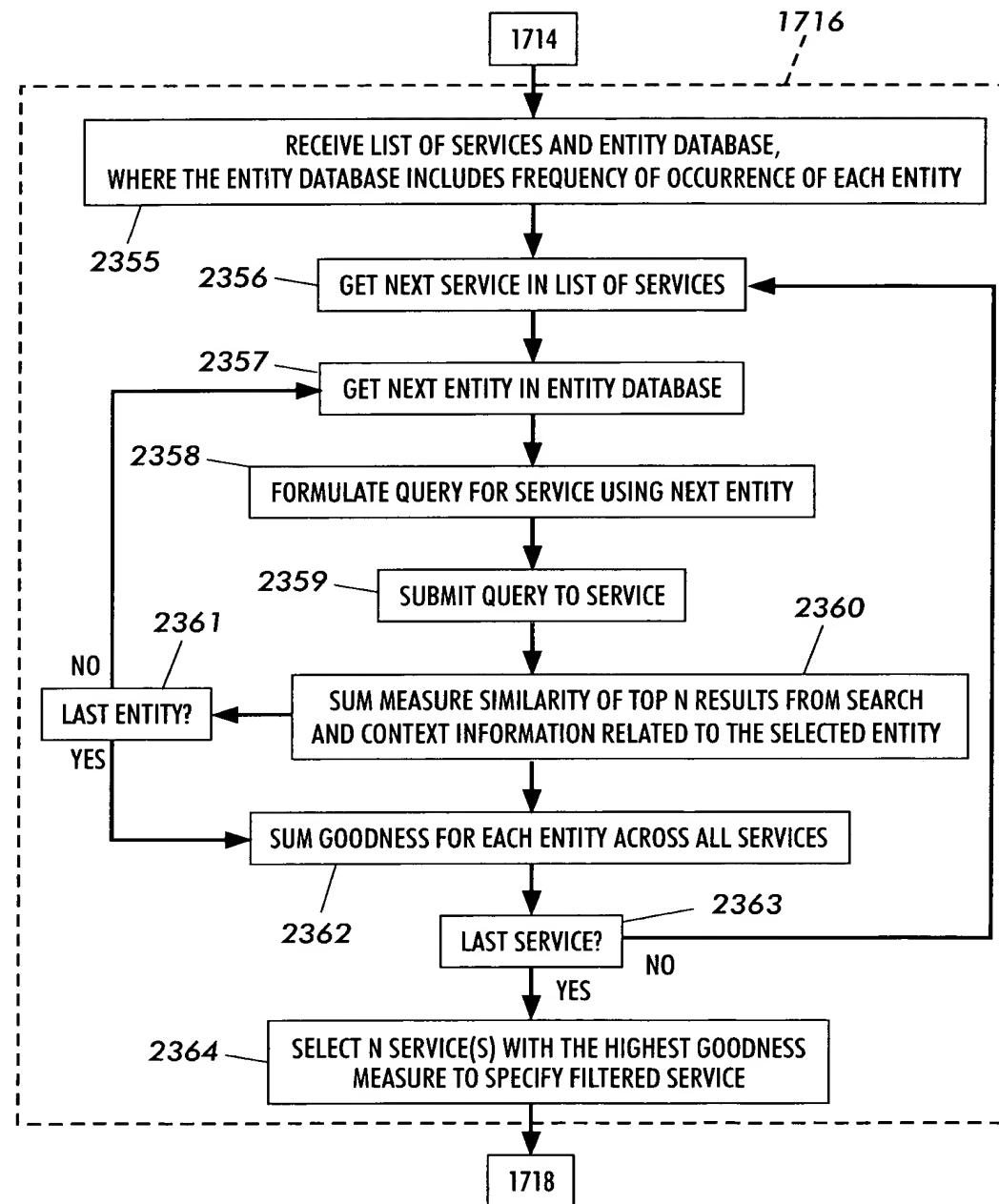
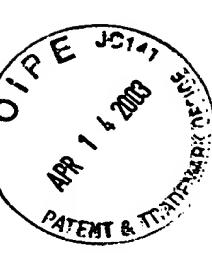


FIG. 23



2176  
09/683,240

24/69

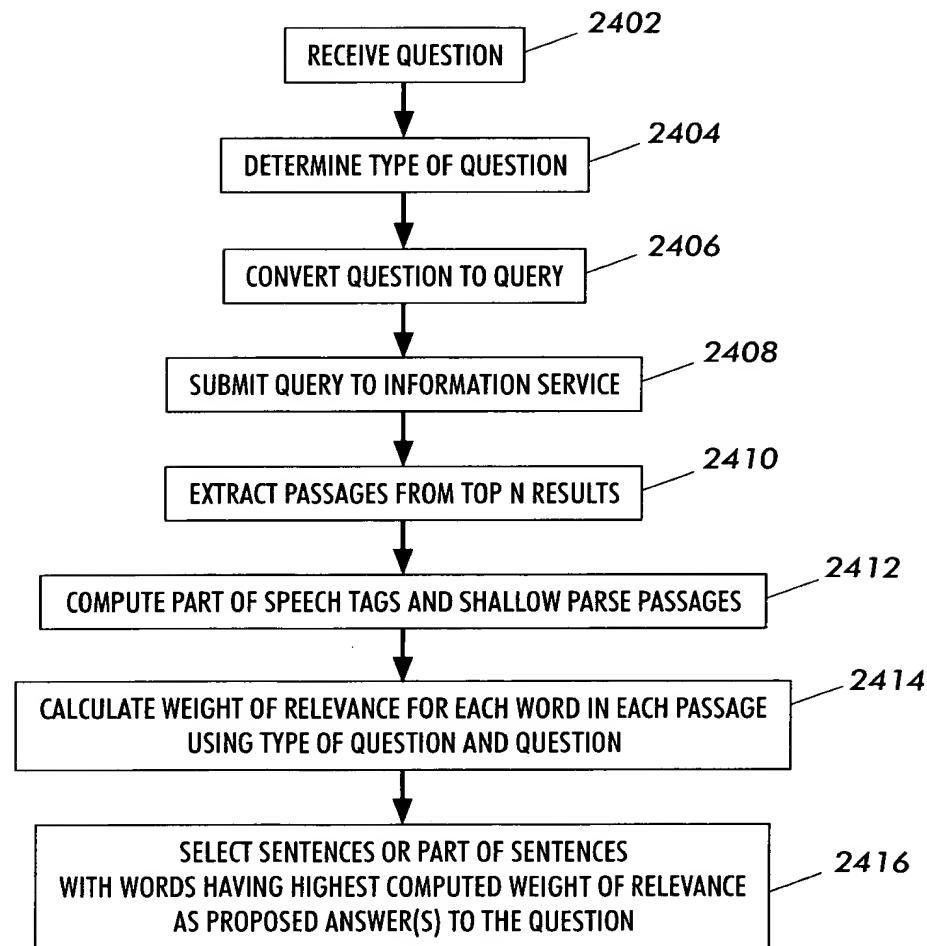


FIG. 24



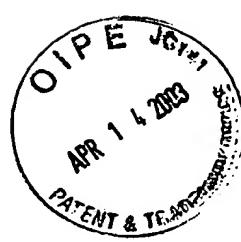
2176  
09/683,240

25/69

E-LEARNING SERVICES ASSOCIATED WITH E-LEARNING PERSONALITY	
SEL	
2504	<input checked="" type="checkbox"/> LINK WORDS OR MULTWORD EXPRESSIONS FOUND IN THE DOCUMENT TO THEIR DEFINITIONS
2506	<input type="checkbox"/> LINK WORDS OR MULTWORD EXPRESSIONS FOUND IN THE DOCUMENT TO THEIR TRANSLATIONS
2508	<input type="checkbox"/> LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO A TUTORIAL CONCERNING EACH TEXT UNIT
2510	<input type="checkbox"/> LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO AN EXPERT CONCERNING IDENTIFIED SUBJECTS
2512	<input type="checkbox"/> LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO AN INTERACTIVE COURSE ABOUT IDENTIFIED SUBJECTS
2514	<input type="checkbox"/> LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO A LIST OF REAL WORLD CLASSES ABOUT IDENTIFIED SUBJECTS
2516	<input checked="" type="checkbox"/> LINK TEXT UNITS TO (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) LEARNING RESOURCES CONCERNING EACH TEXT UNITS

2502

FIG. 25



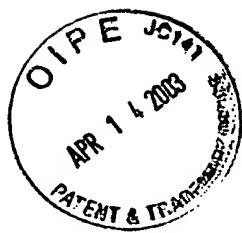
2176  
09/683,240

26/69

LANGUAGE LEARNING SERVICES ASSOCIATED WITH LANGUAGE LEARNING PERSONALITY	
SEL	
2604	<input checked="" type="checkbox"/> LINK WORDS OR MULTWORD EXPRESSIONS FOUND IN THE DOCUMENT TO THEIR DEFINITIONS
2606	<input type="checkbox"/> LINK WORDS OR MULTWORD EXPRESSIONS FOUND IN THE DOCUMENT TO THEIR TRANSLATIONS
2608	<input type="checkbox"/> LINK EACH SENTENCE, OR PHRASE, TO A GRAMMATICAL DESCRIPTION OF THE STRUCTURE OF THE SENTENCE OR PHRASE
2610	<input type="checkbox"/> LINK EACH WORD, MULTWORD EXPRESSION, PHRASE OR SENTENCE TO OTHER INSTANCES OF THE SAME IN DIFFERENT CONTEXTS FROM THE PRESENT
2612	<input type="checkbox"/> LINK EACH WORD, MULTWORD EXPRESSION, PHRASE OR SENTENCE TO ONE OR MORE INTERACTIVE GRAMMAR EXERCISES CONCERNING THAT ELEMENT
2614	<input checked="" type="checkbox"/> LINK EACH TEXT UNIT (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) TO A LIST OF REAL WORLD CLASSES ABOUT IDENTIFIED LANGUAGE
2616	<input checked="" type="checkbox"/> LINK TEXT UNITS TO (I.E., DOCUMENT, PARAGRAPH, PHRASE, WORD) LANGUAGE LEARNING RESOURCES CONCERNING EACH TEXT UNITS

2602

FIG. 26



2176  
09/683,240

27/69

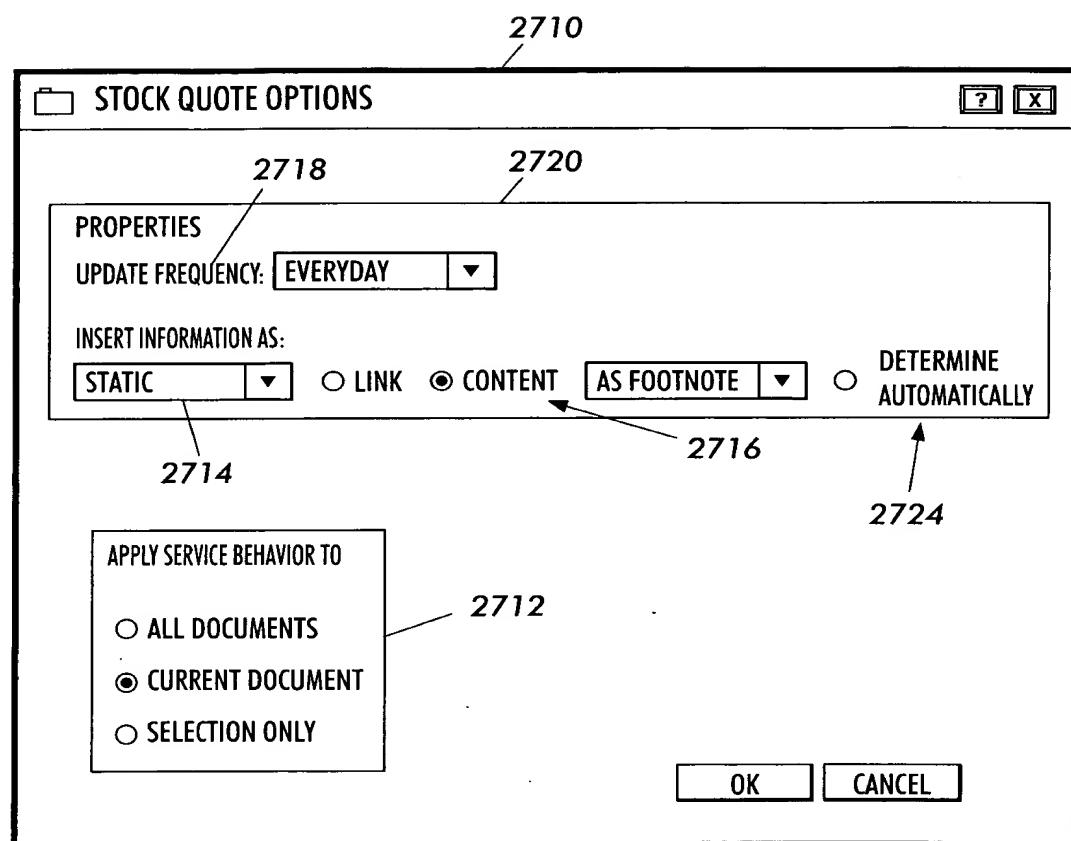
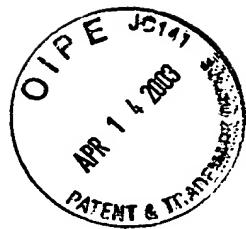


FIG. 27



2176  
09/683,240

28/69

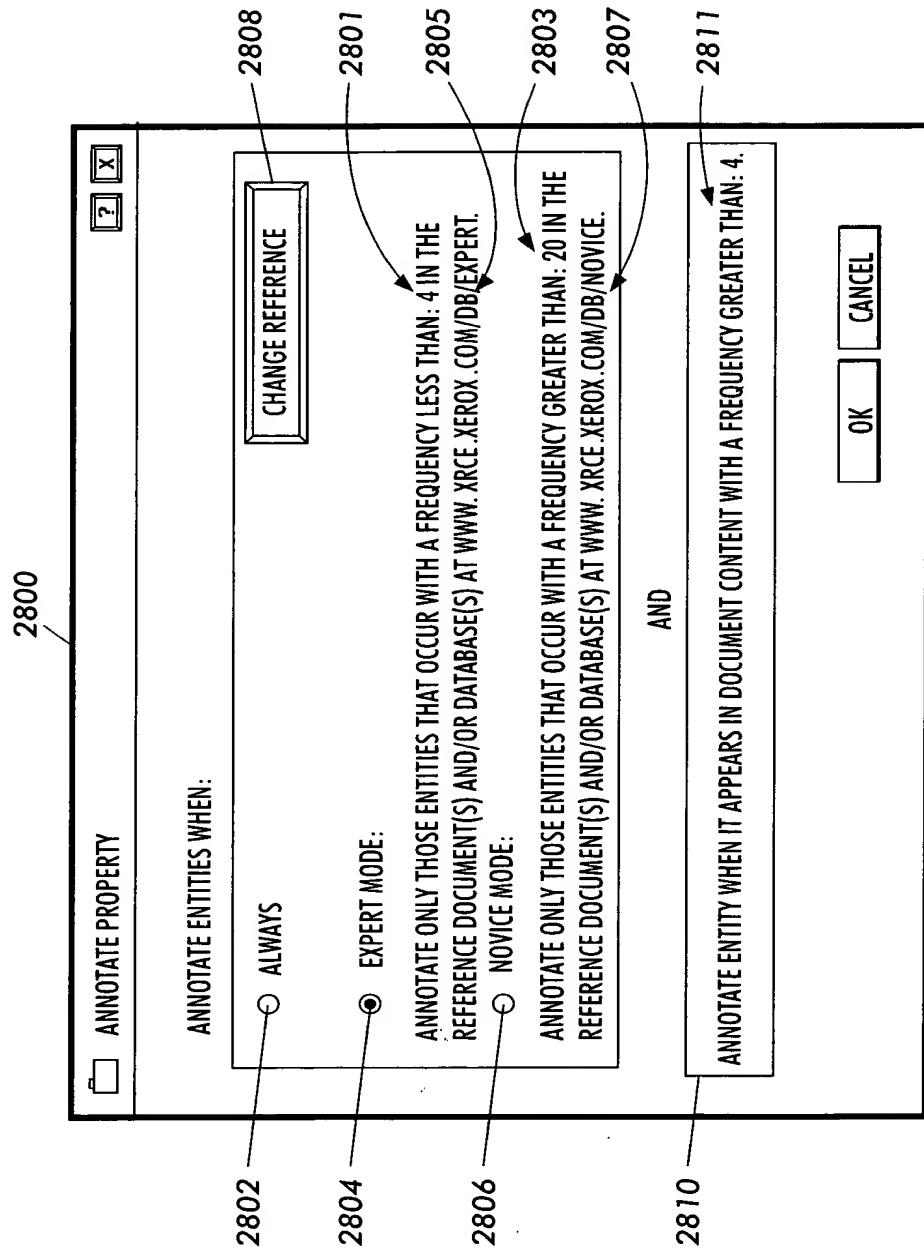


FIG. 28

29/69

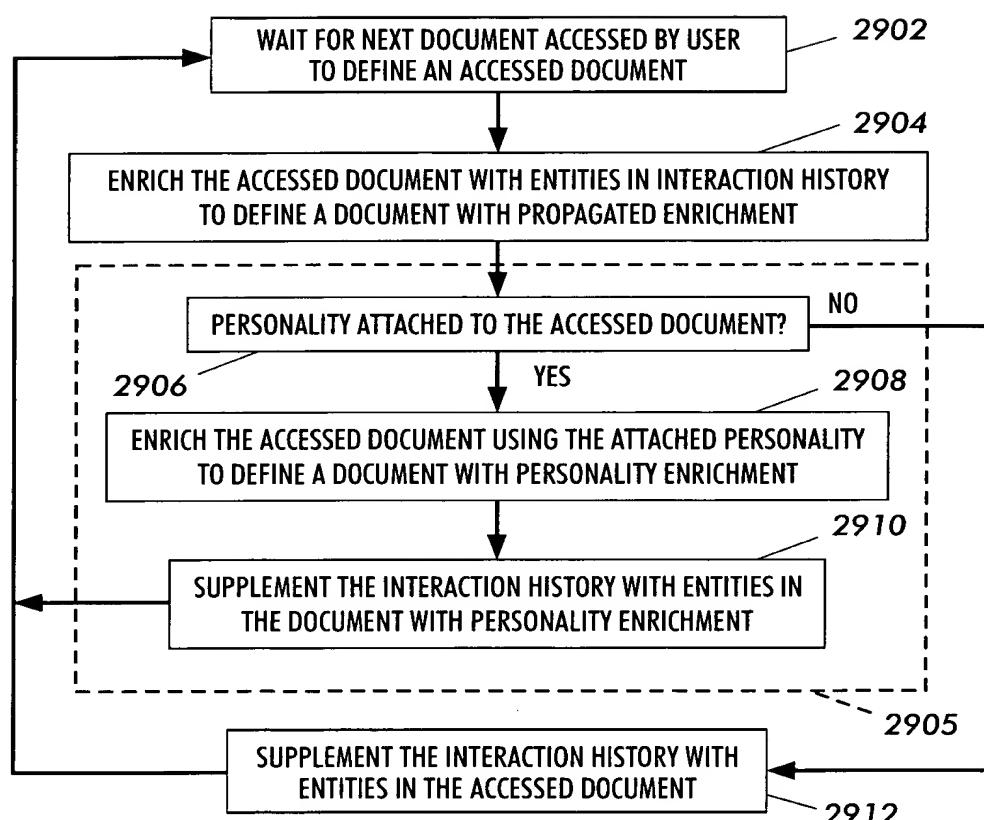
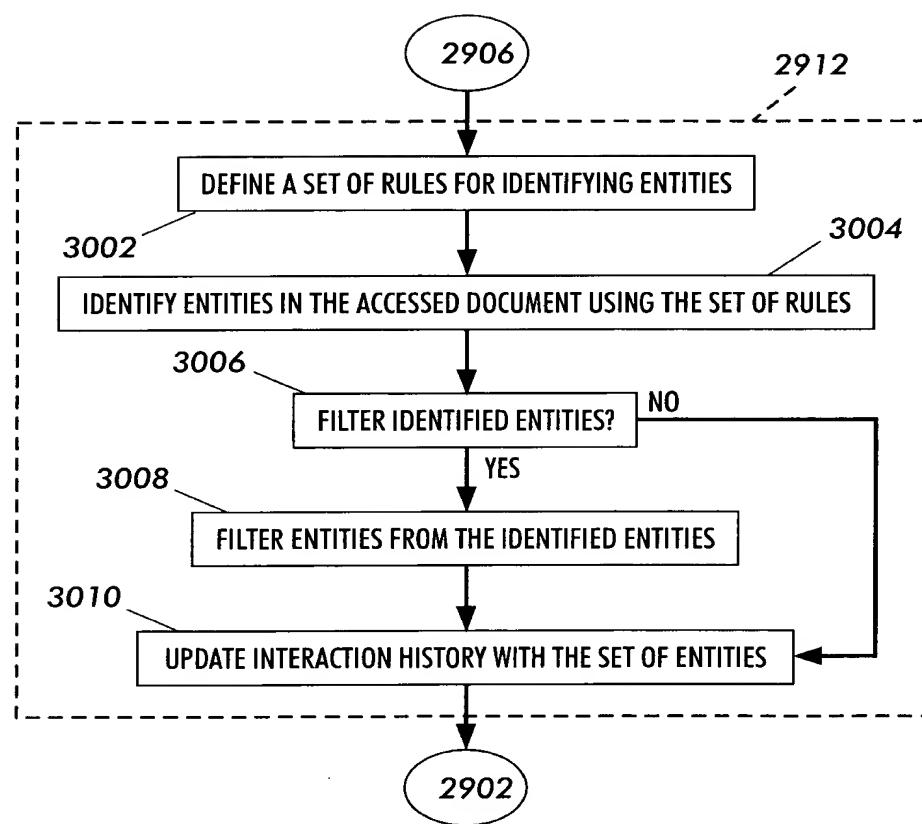


FIG. 29



**FIG. 30**

31/69

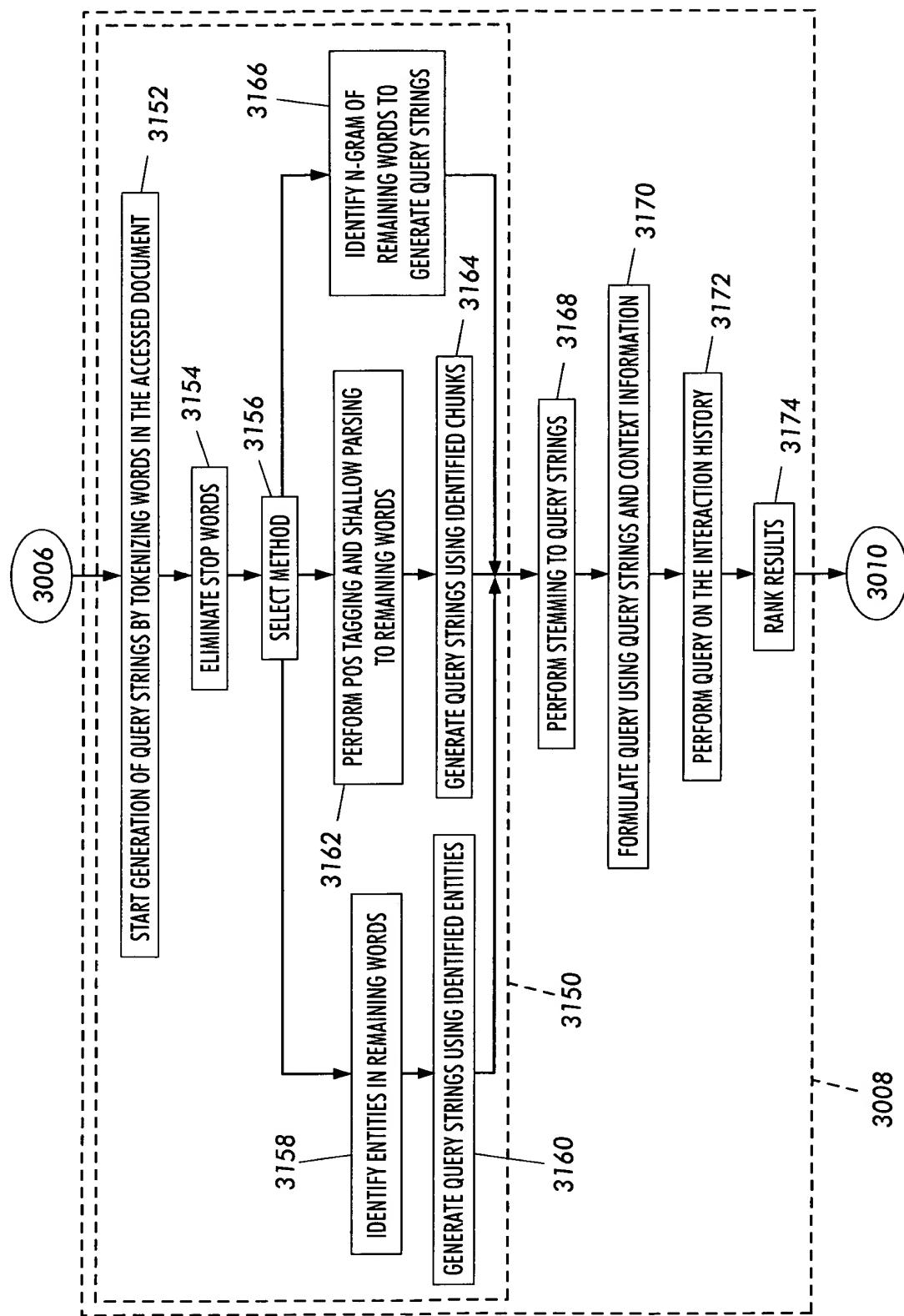
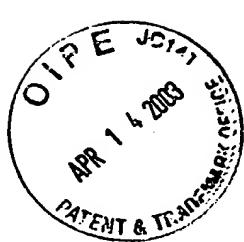


FIG. 31

3008



32/69

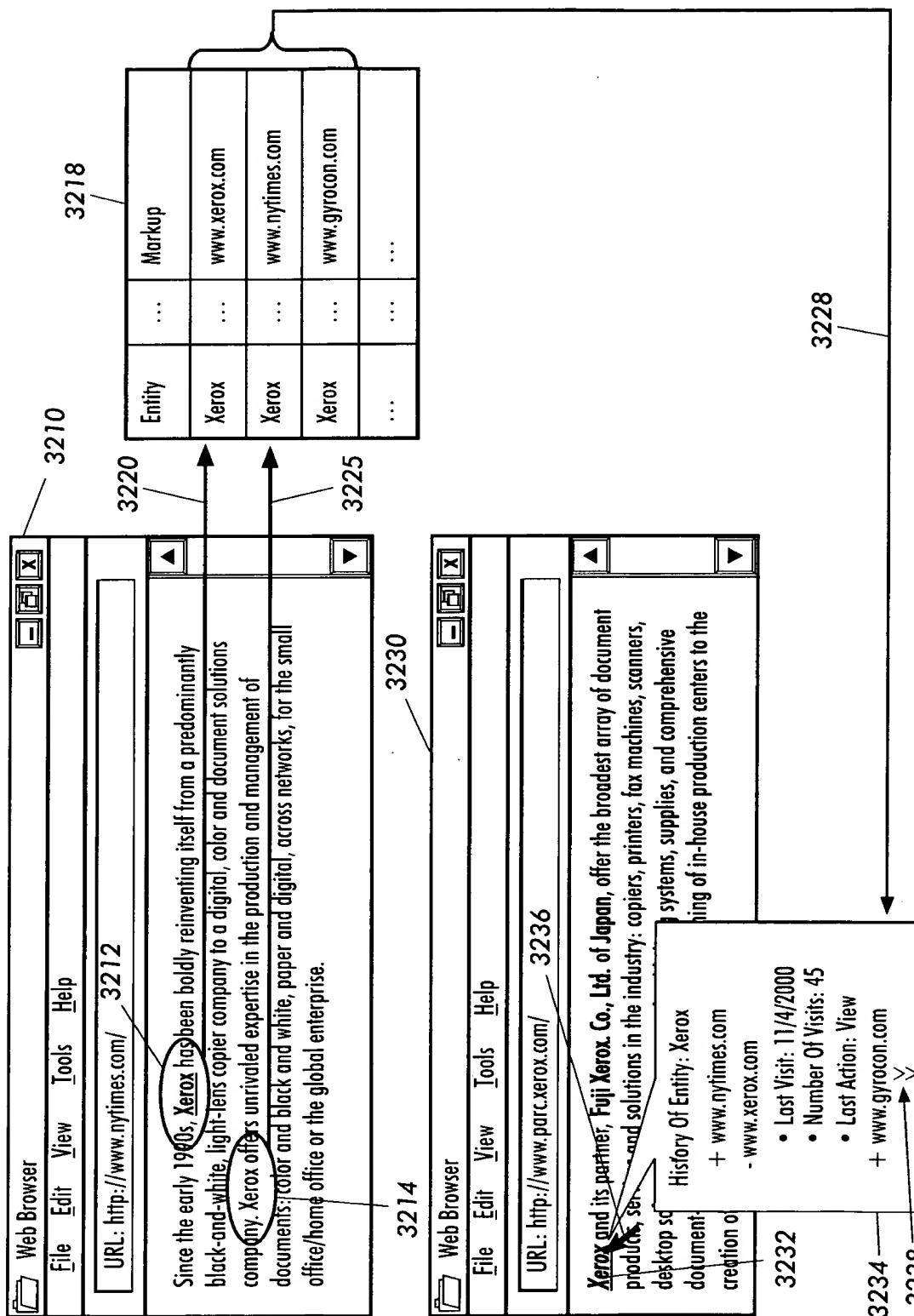
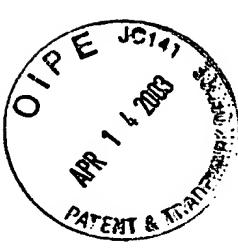


FIG. 32



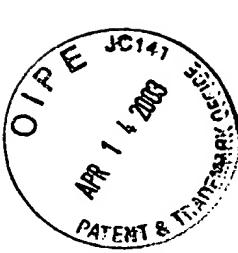
2176  
09/683,240

33/69

3218

ENTITY	ENTITY TYPE	VISITS	LAST VISIT	ACTION LIST	DMOZ CATEGORY	LOCATION	POS	MARKUP
XEROX	BUSINESS	45	11/4/2000	VIEW(13), PRINT(2)	BUSINESS	40	NOUN	WWW.XEROX.COM
XEROX	BUSINESS	1	5/1/2001	VIEW(1)	FINANCE	689	NOUN	WWW.NYTIES.COM
XEROX	BUSINESS	4	2/22/2001	EMAIL(1), VIEW(3)	COMPUTER	343	NOUN	WWW.GYROCON.COM
...	...	...	...	...	...	...	...	...

FIG. 33



34/69

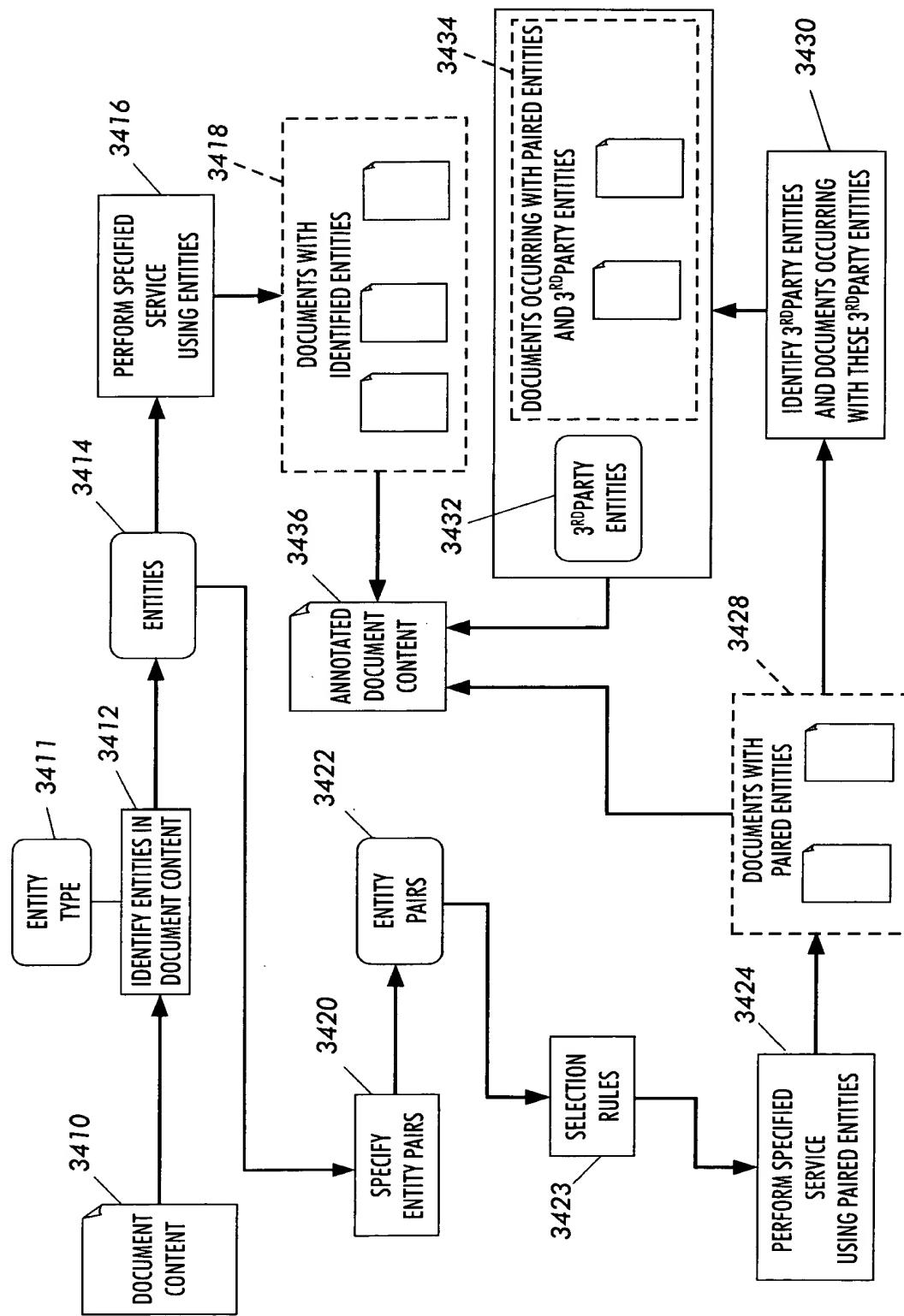


FIG. 34

35/69

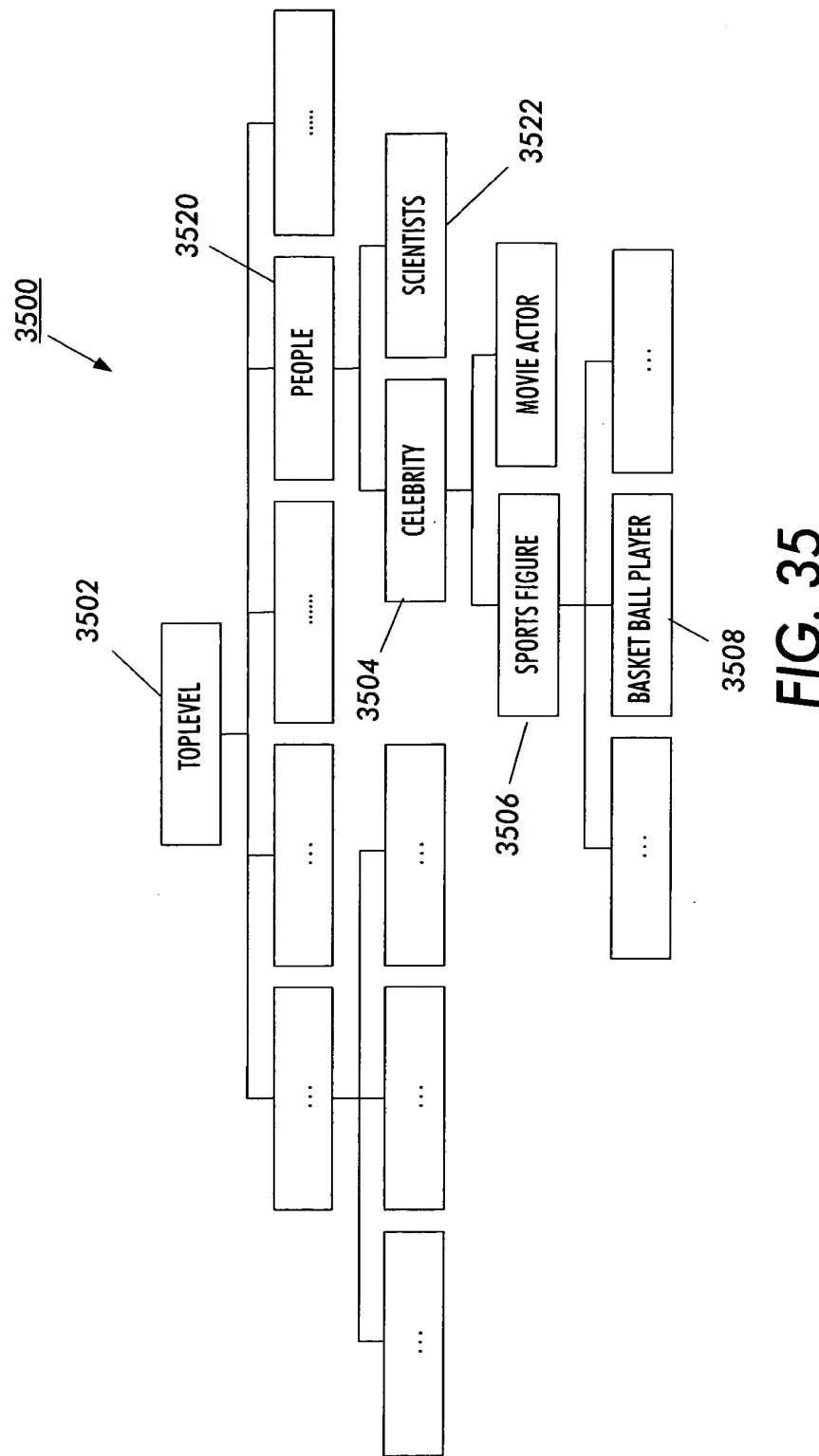


FIG. 35

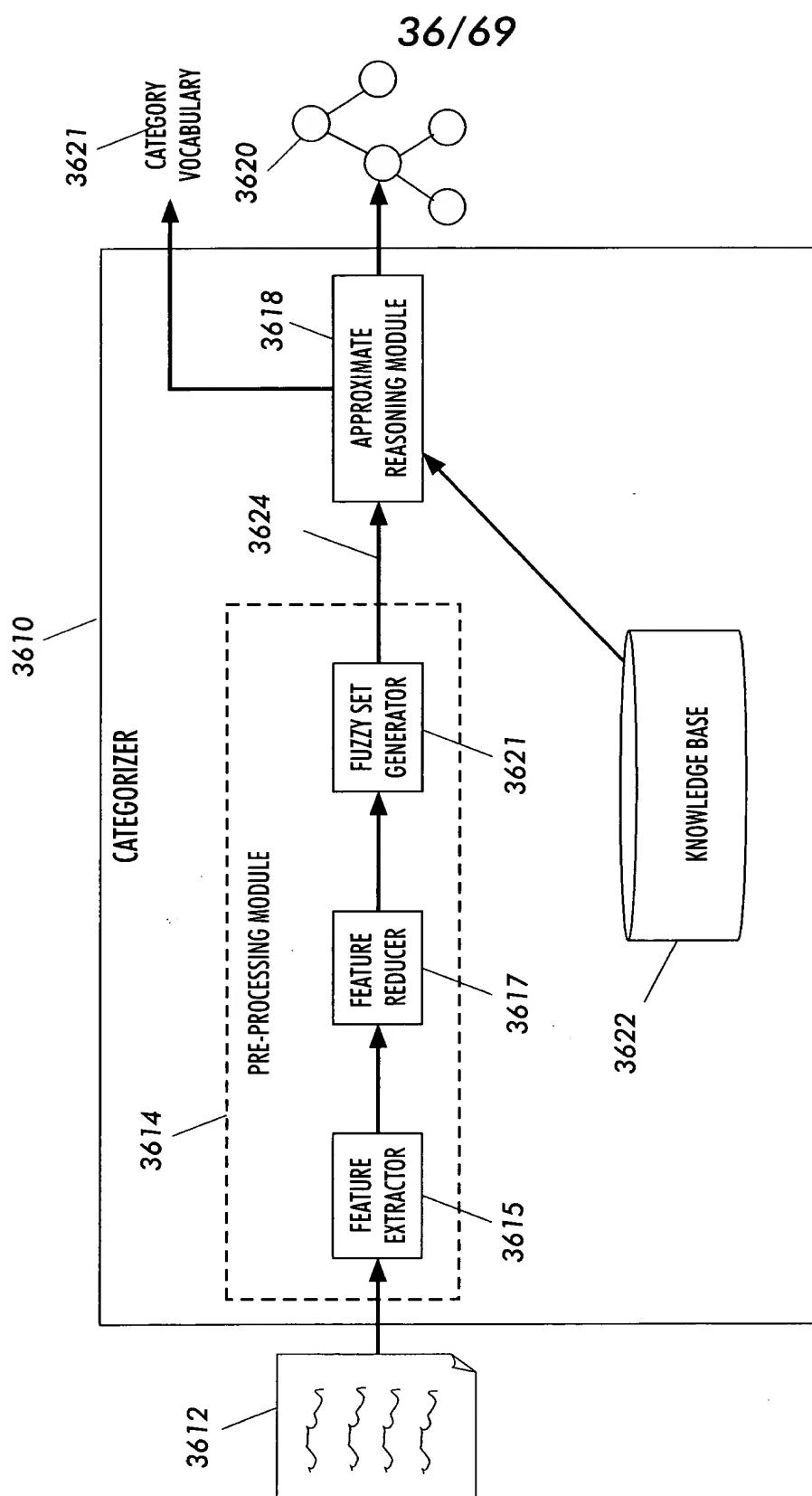
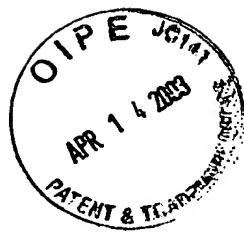


FIG. 36



2176  
09/683,240

37/69

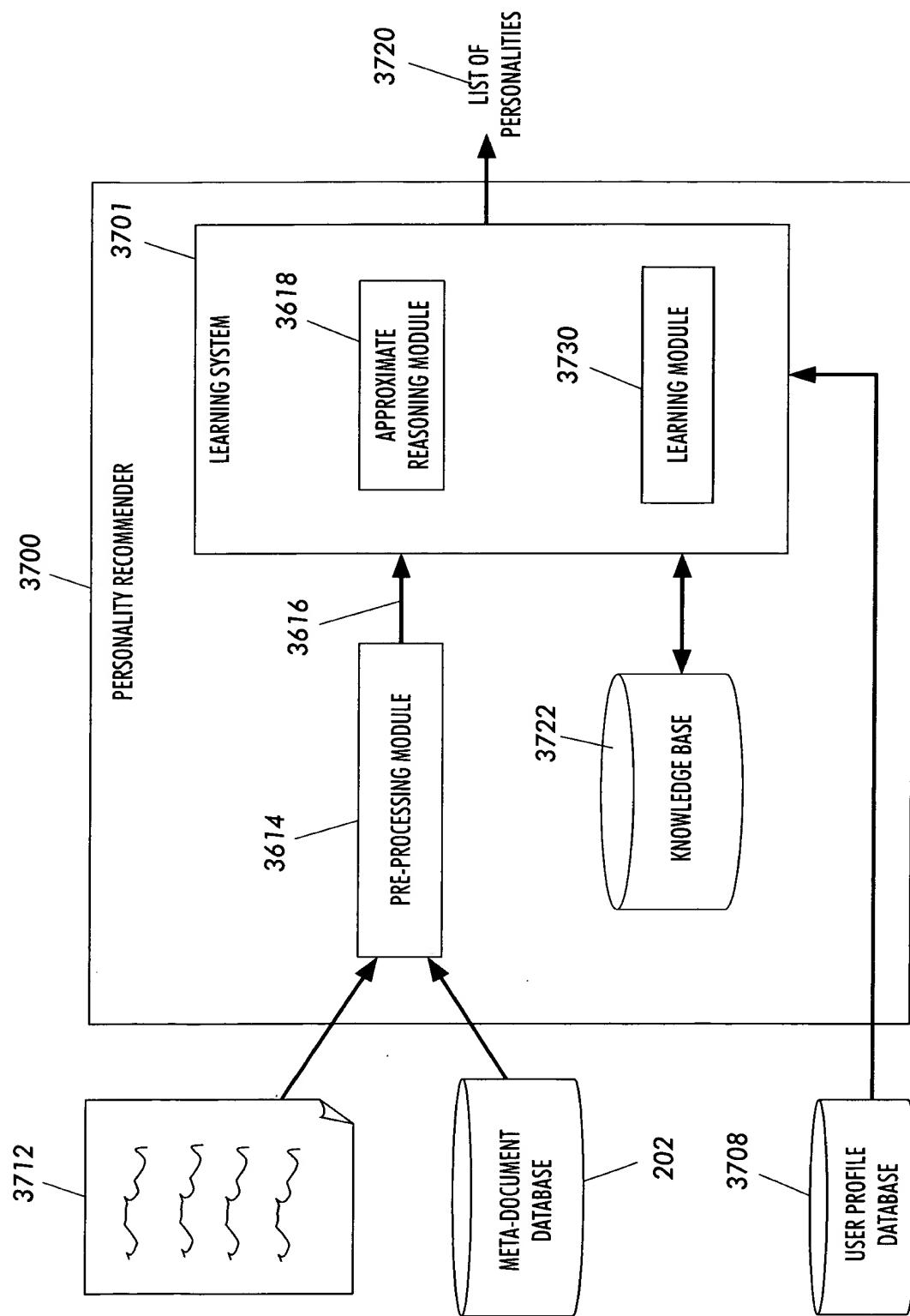


FIG. 37

38/69

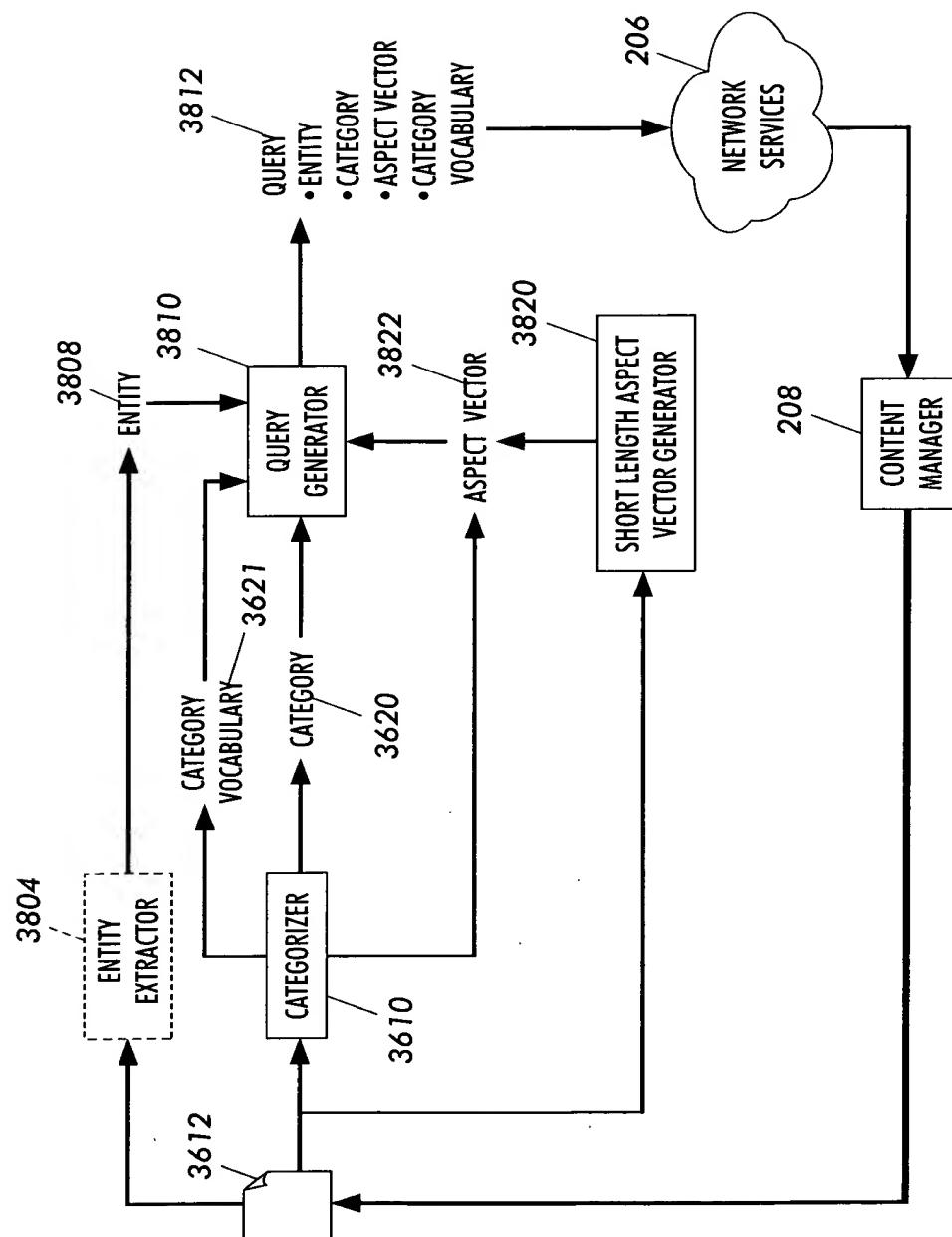
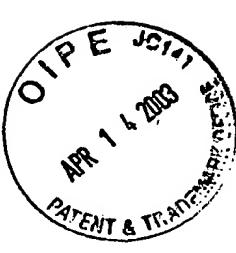


FIG. 38



2176  
09/683,240

39/69

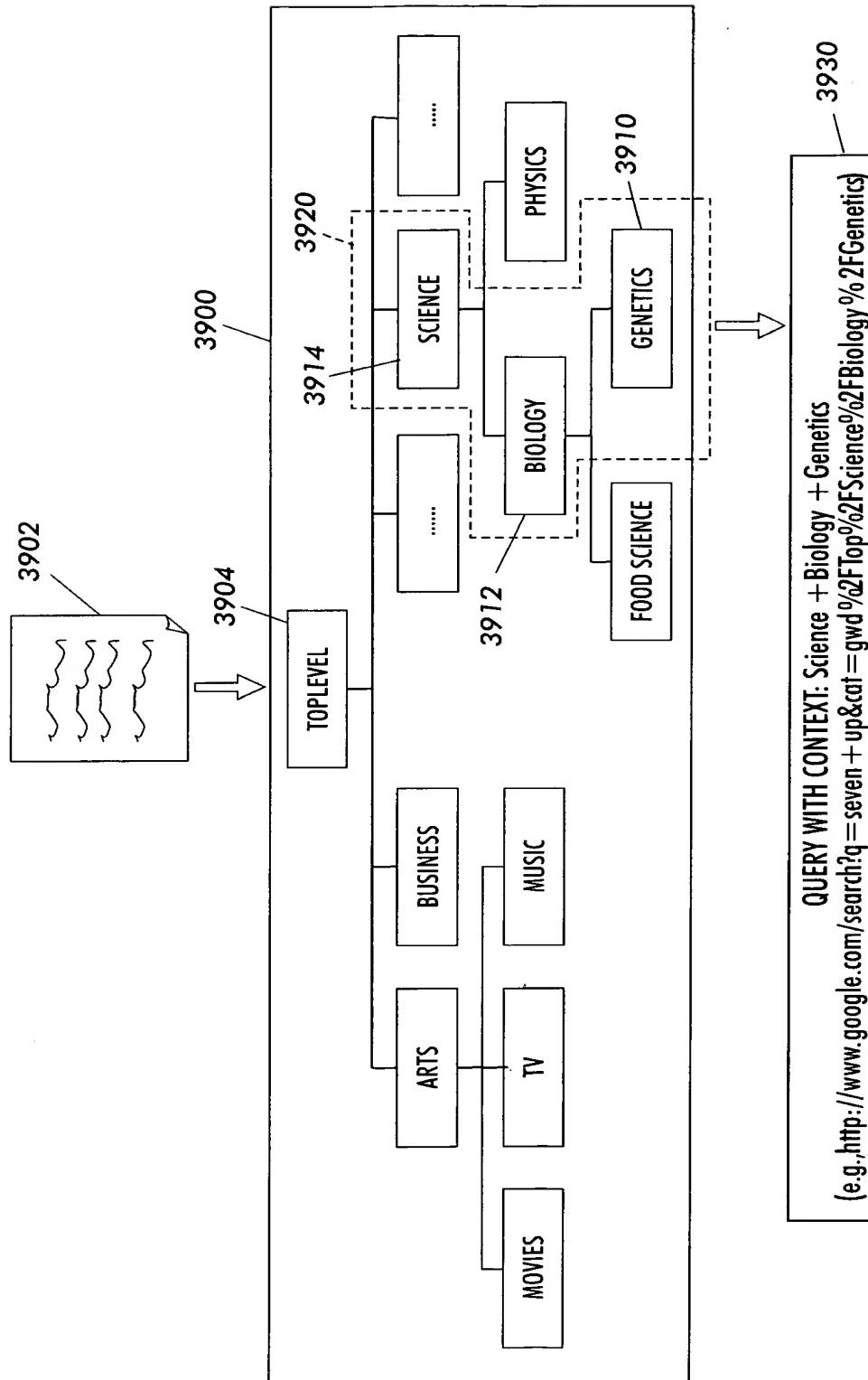


FIG. 39



2176  
09/683,240

40/69

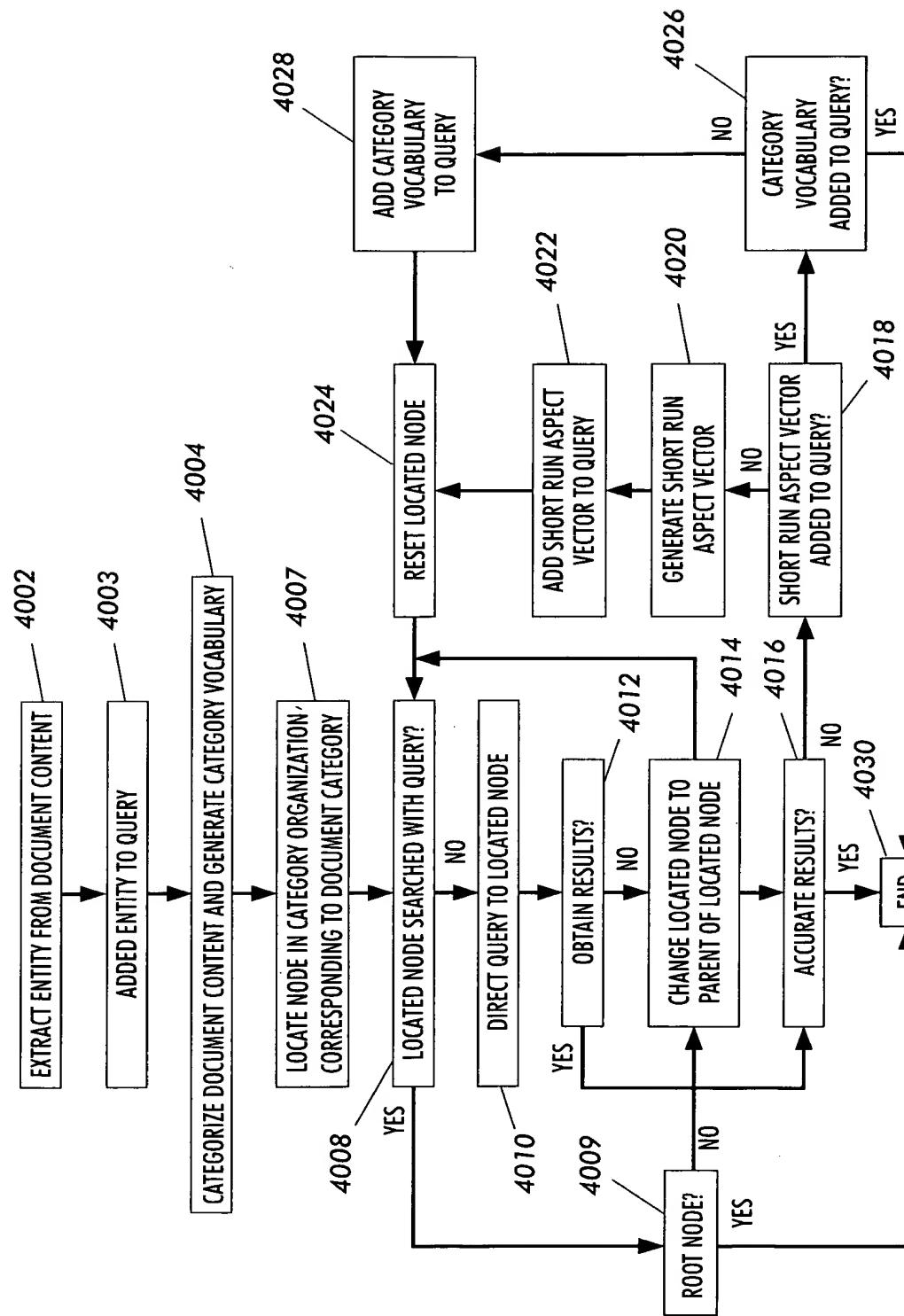


FIG. 40



41/69

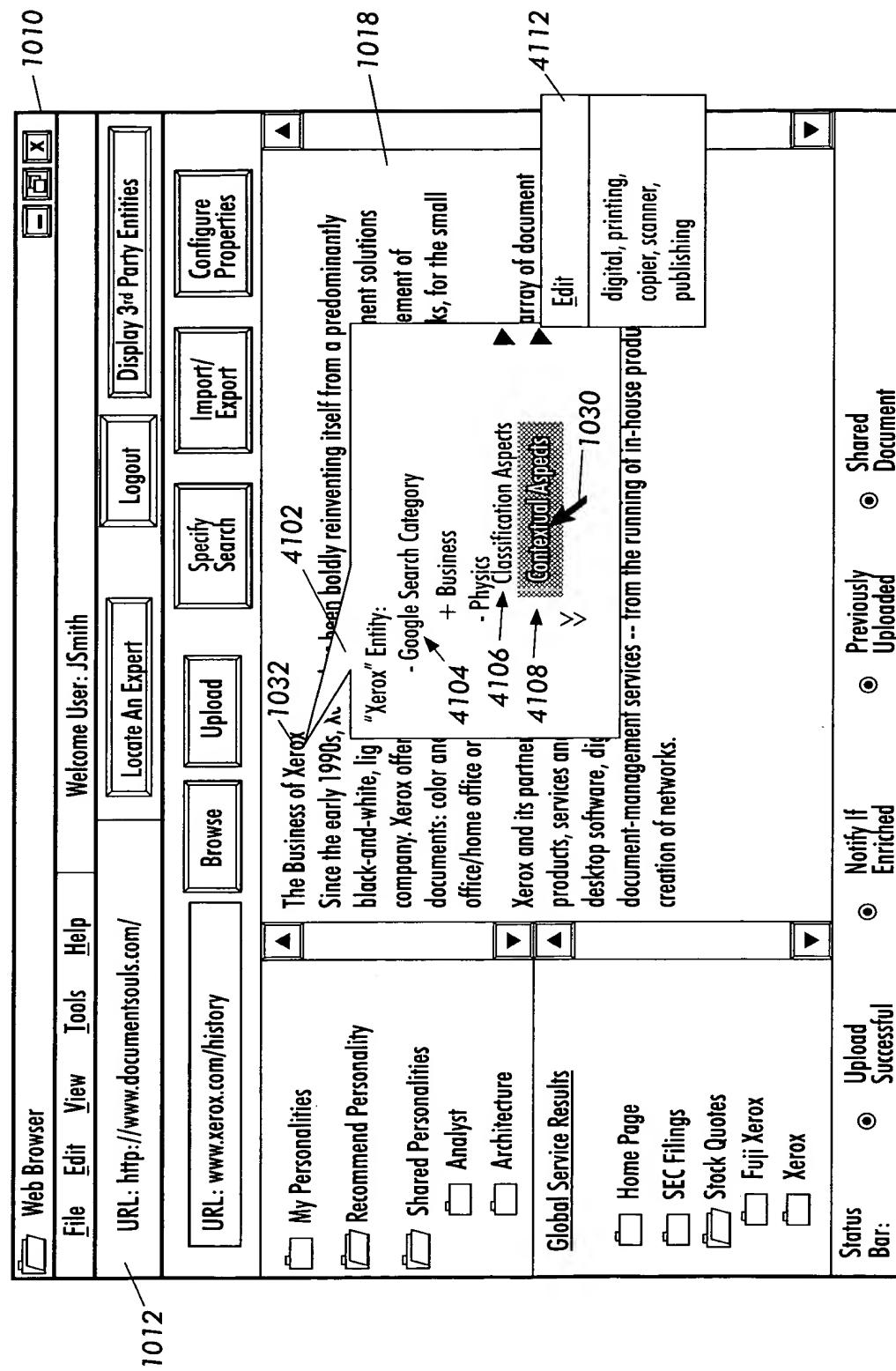


FIG. 41

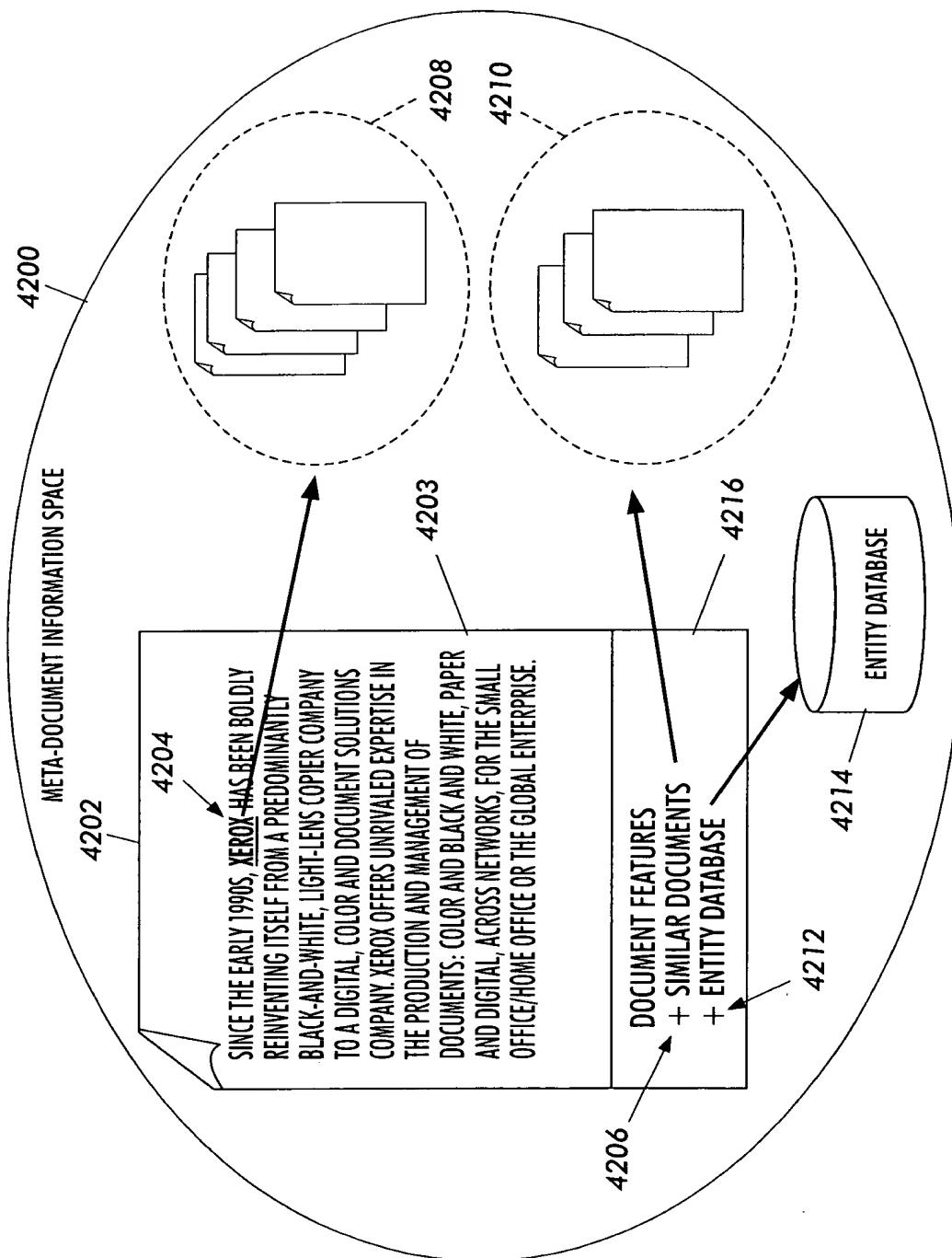


FIG. 42

43/69

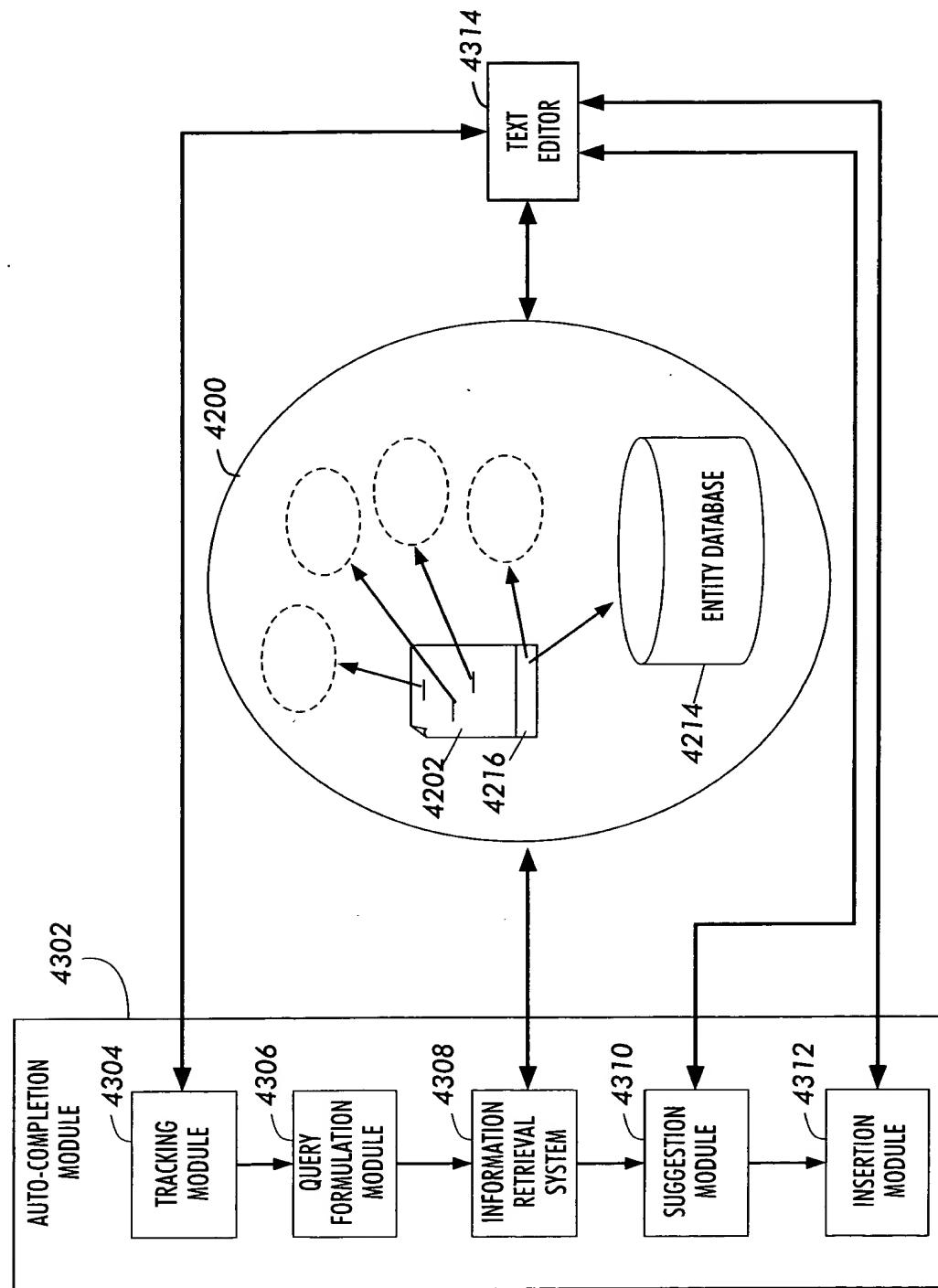


FIG. 43

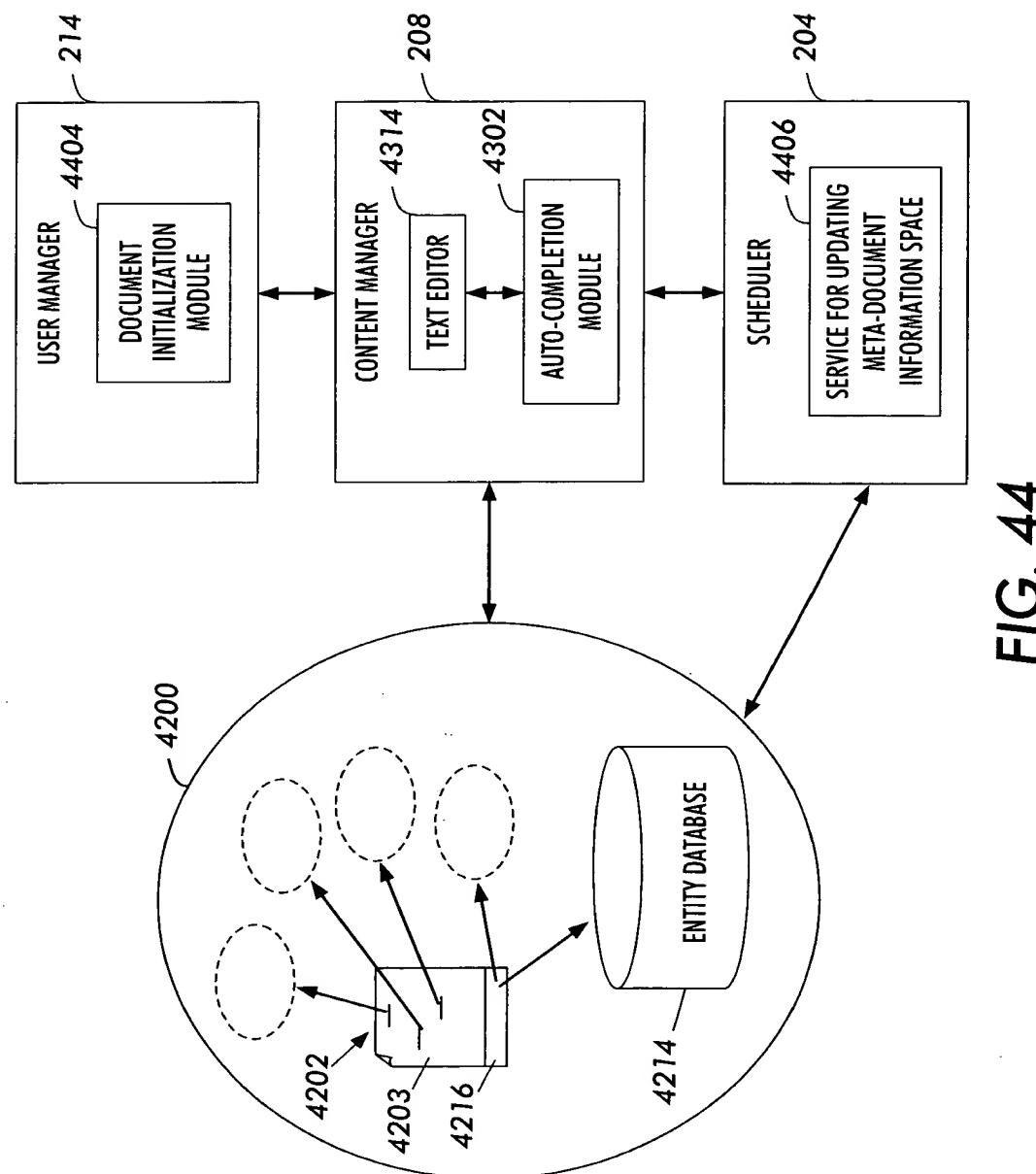


FIG. 44

45/67

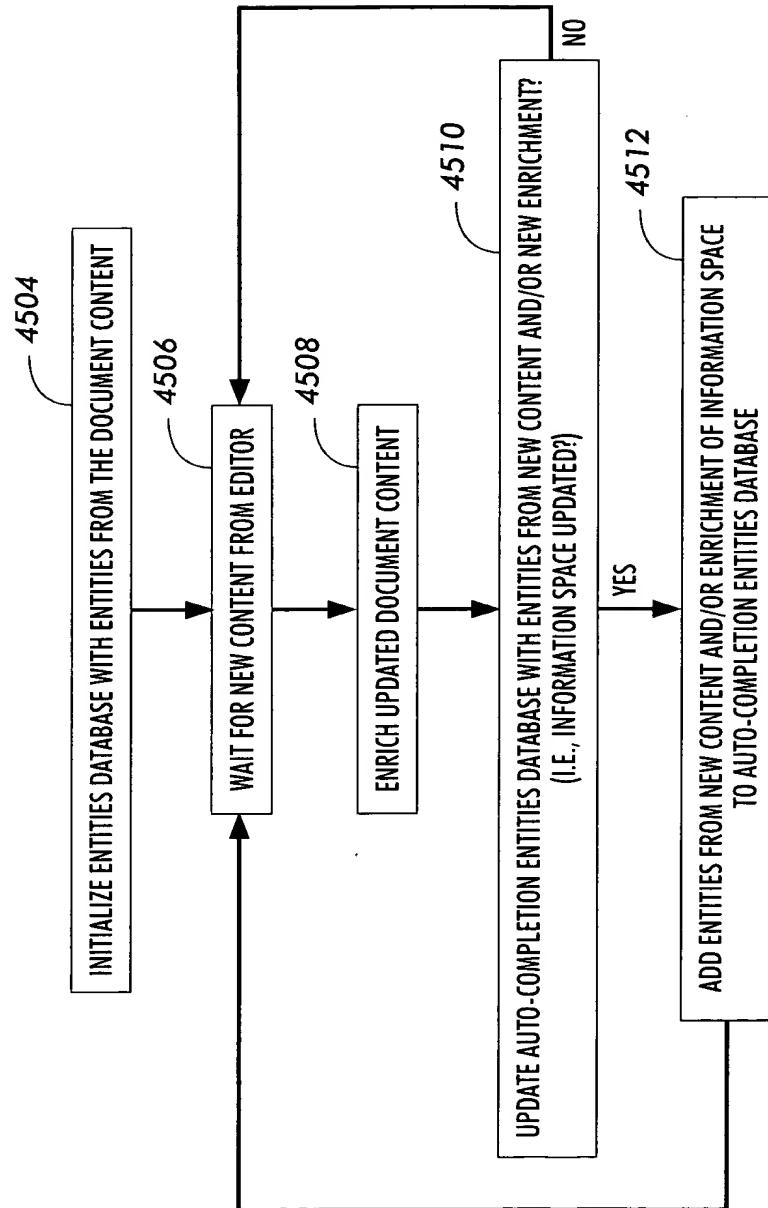


FIG. 45

46/67

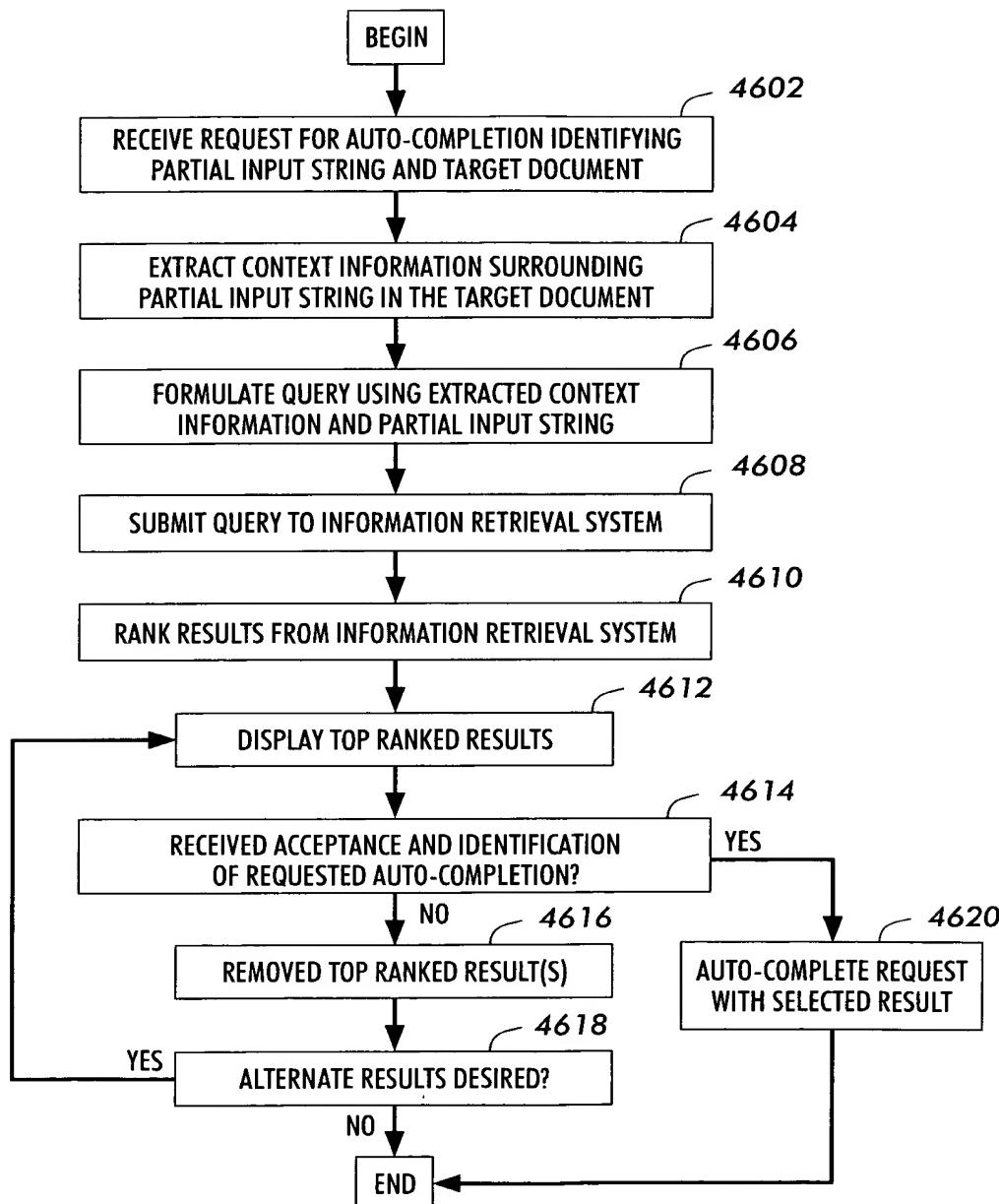
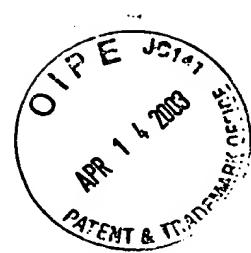


FIG. 46



2176  
09/683,240

47/69

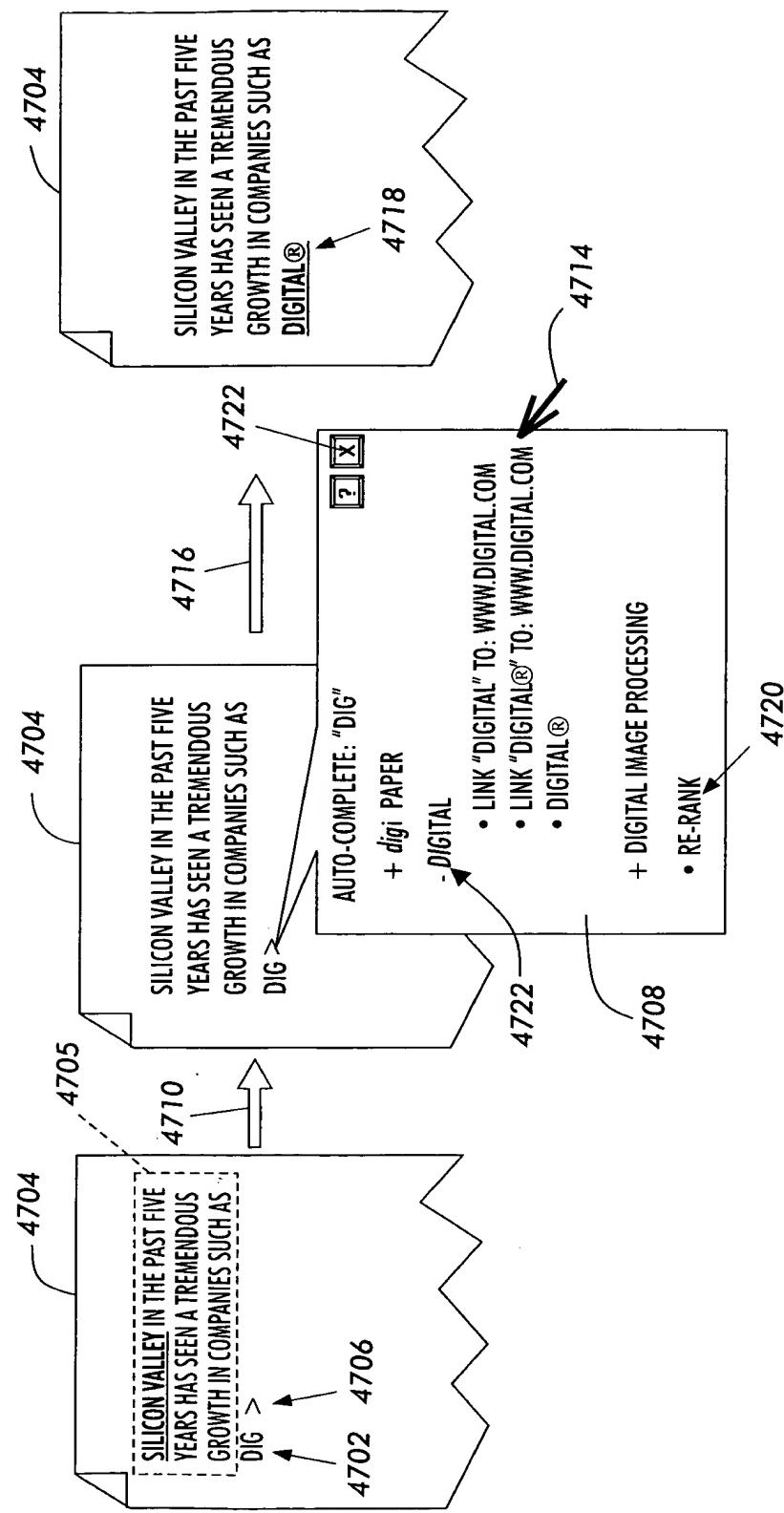


FIG. 47



2176  
09/683,240

48/69

(A) KEY	(B) ENTITY	(C) MARKUP	(D) FORMAT	(E) ORIGIN CONTENT	(F) ORIGIN POSITION	(G) PART OF SPEECH	(H) CONTEXT
Sili	Silicon Valley	www.siliconvalley.com	Capitalize	TARGET DOCUMENT 1	Noun	Science	
Dig	Digital	www.digital.com	®	www.digital.com	434	Adjective	Science
Dig	DigiPaper	NONE	Bold	www.parc.xerox.com	23	Noun	Science
Dig	digress	NONE	NONE	www.nytimes.com	343	Verb	Politics
...	...	...	...	...	...	...	...

FIG. 48

49/69

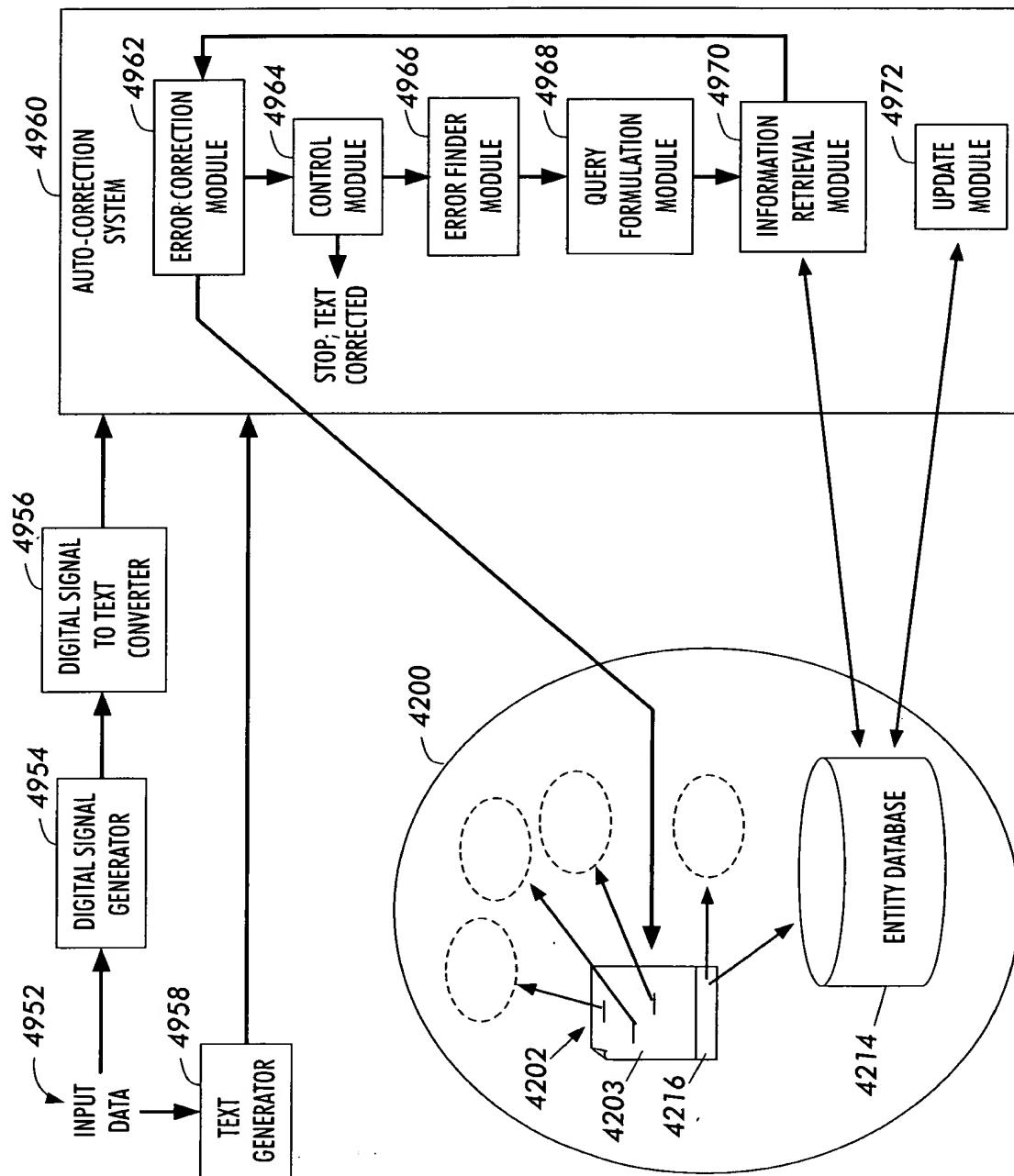


FIG. 49

50/69

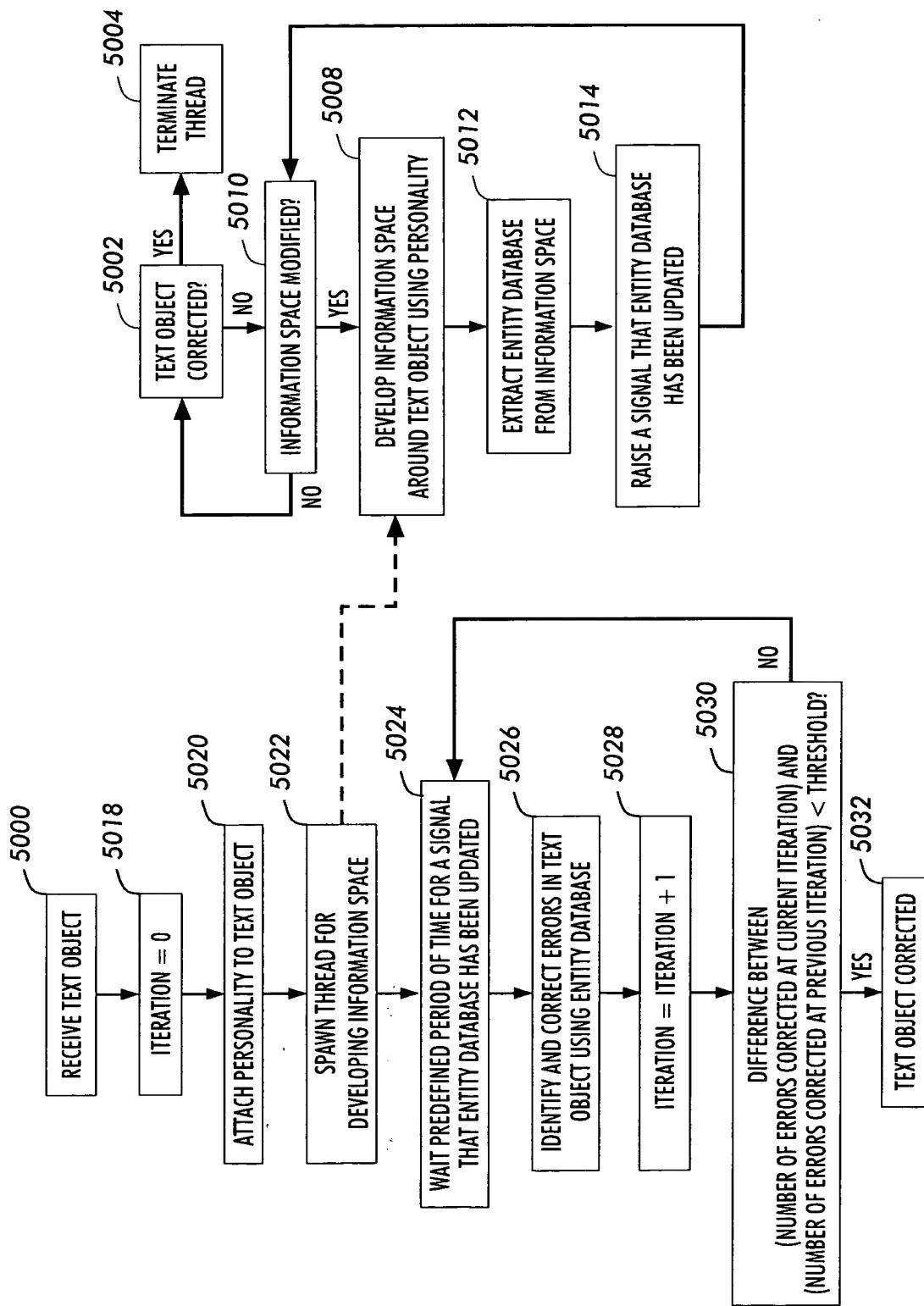


FIG. 50

51/69

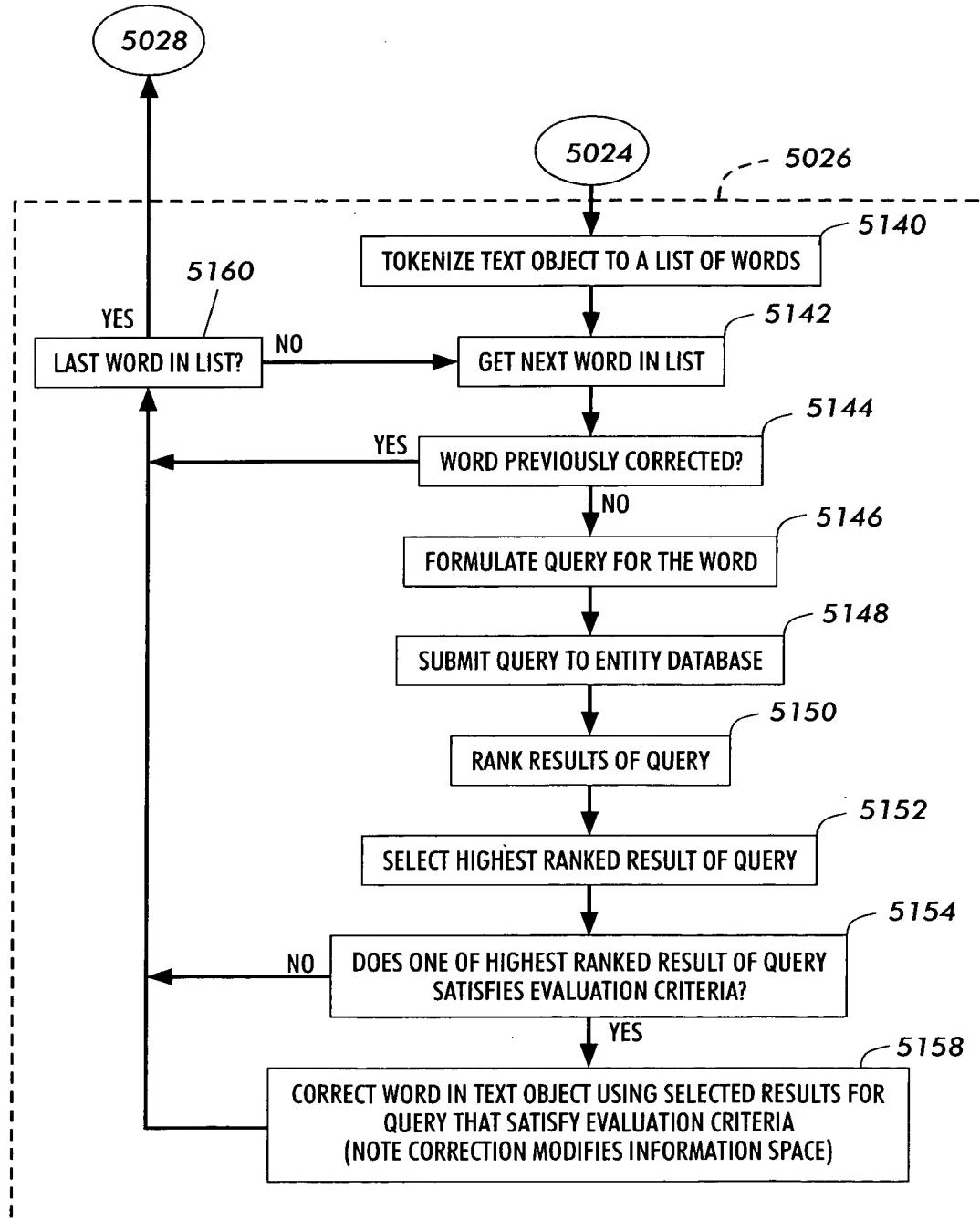


FIG. 51

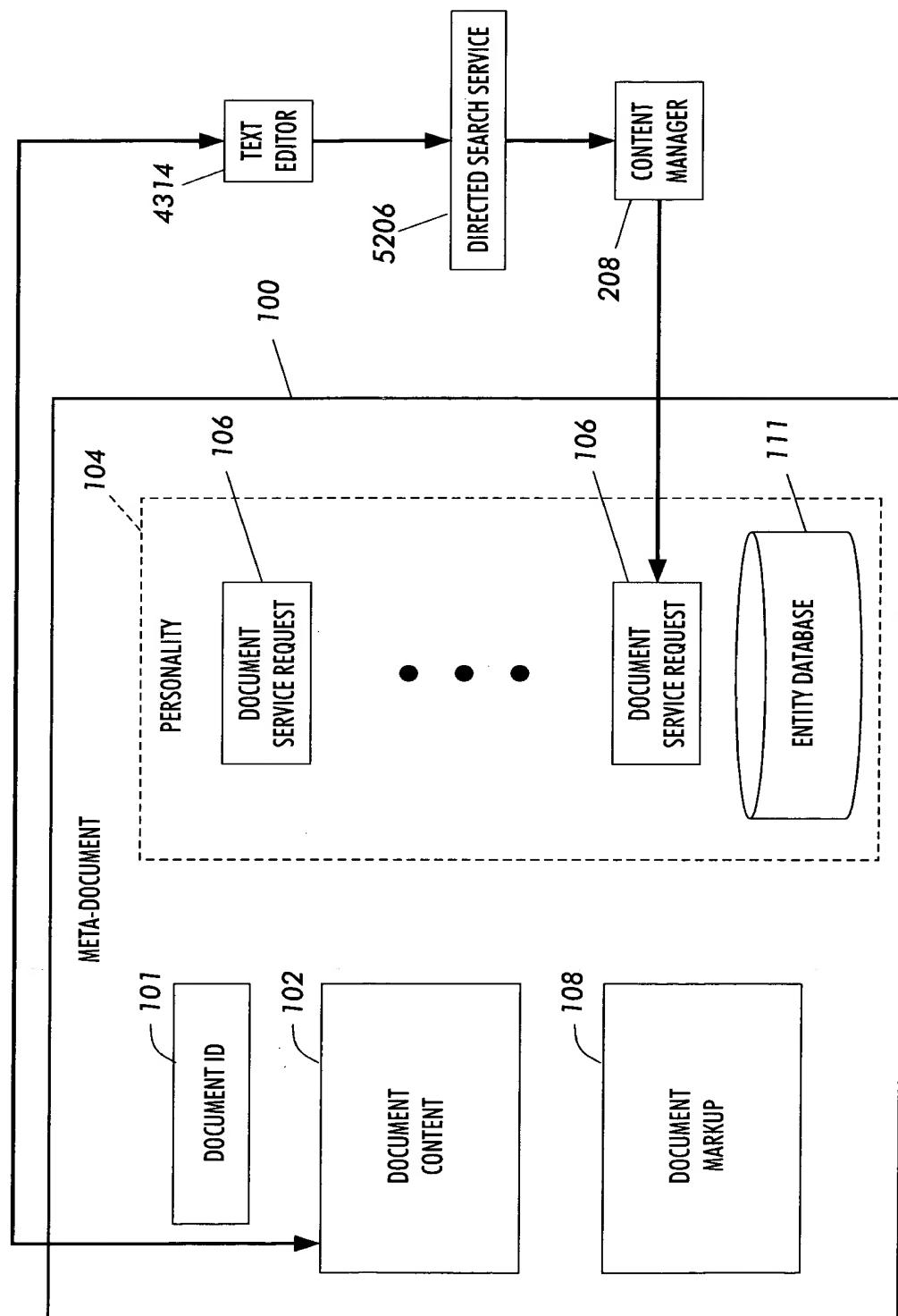
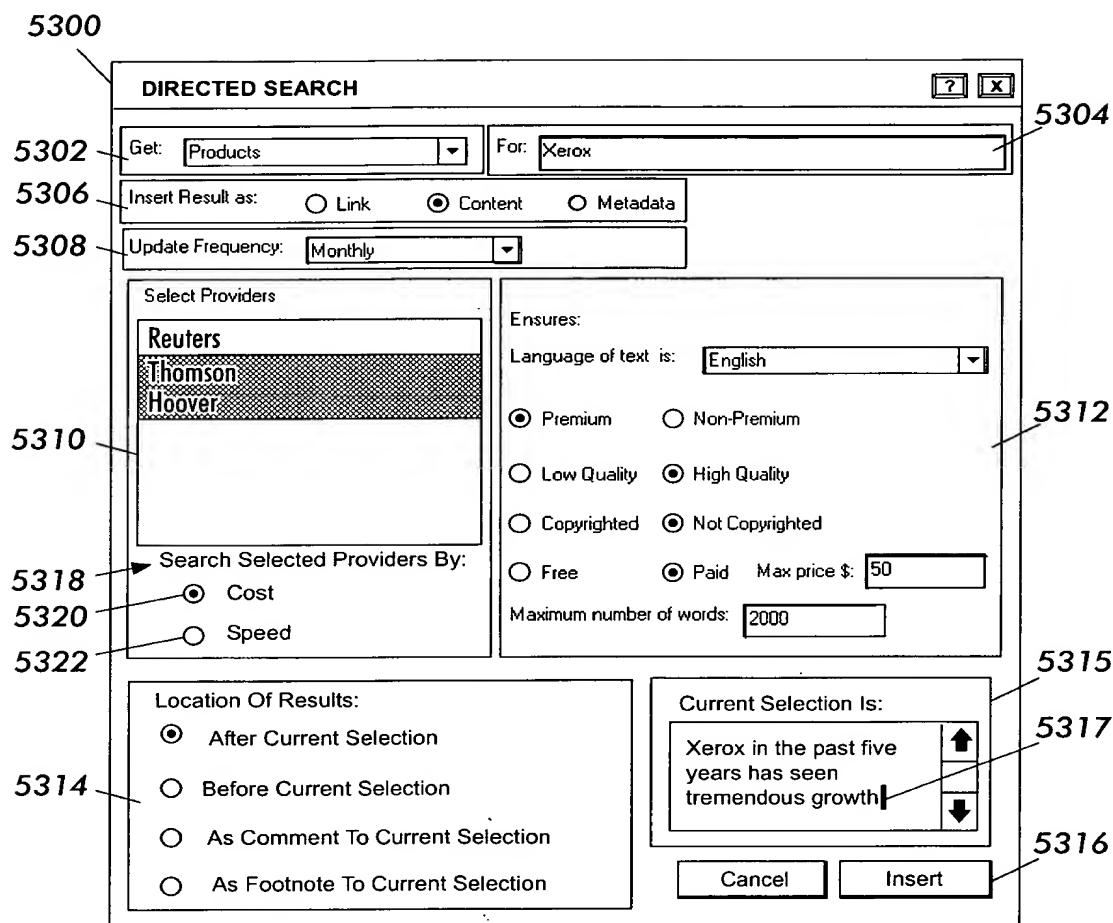
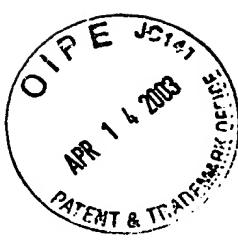


FIG. 52





2176  
09/683,240

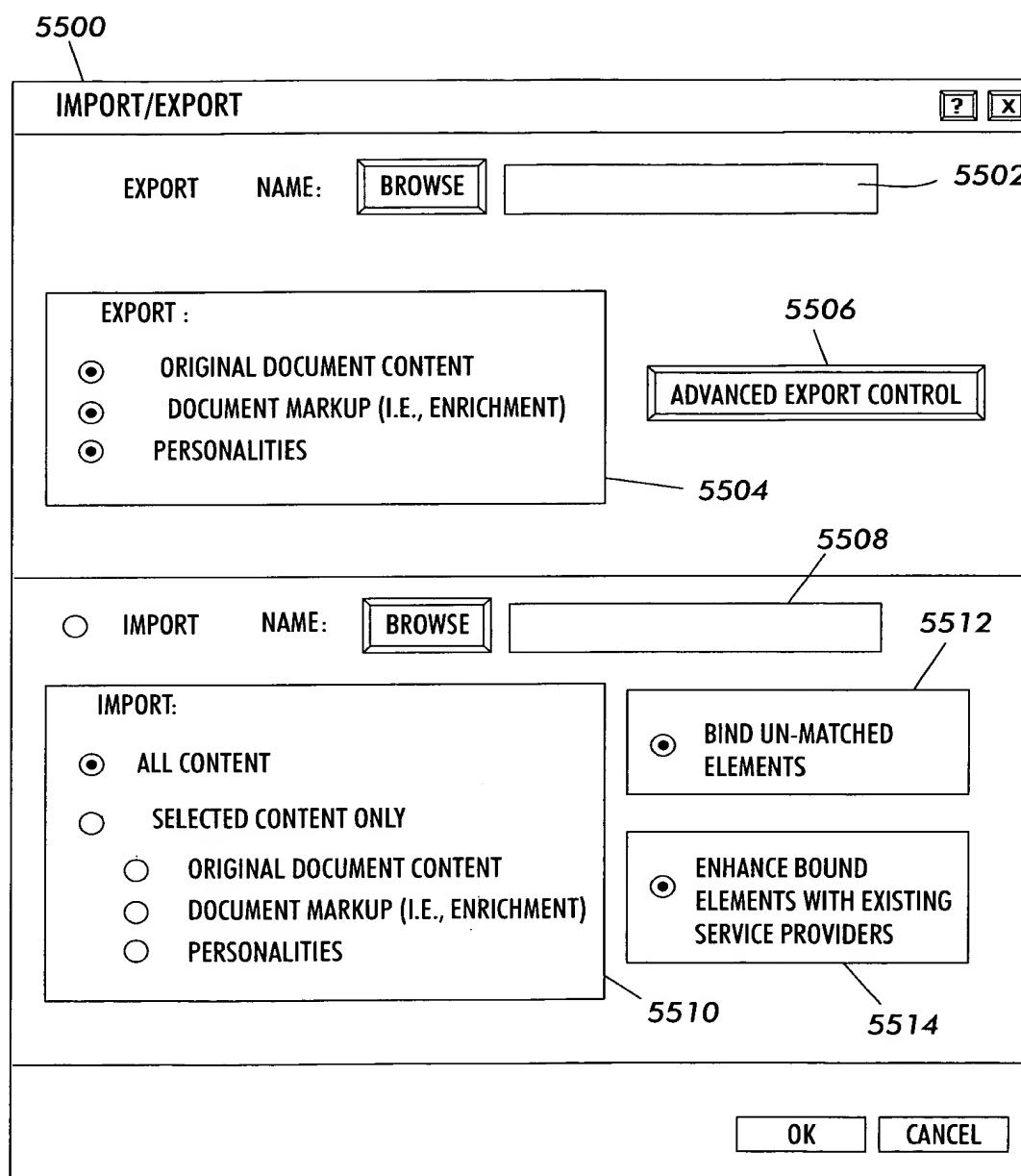
54/69

5400

```
<service id='Products'  
refresh=30  
insert='Content'  
language='En'  
quality='Premium&High&Not-copyrighted&paid'  
maxprice='50'  
maxwords='2000'  
search='cost'>  
<entity>Xerox  
<offsets>13 18 456 461 710 715</offsets>  
<provider name='Thomson' id=18></provider>  
<provider name='Reuters' id=19></provider>  
</entity>  
</service>
```

FIG. 54

55/69



**FIG. 55**

```
<!-- SERVER ID>
<docSouls:serverid> DS v1.0 server name </docSouls:serverid>
  <!-- PROPERTIES OF DOCUMENT>
    <docSouls:docid> xxx </docSouls:docid>
      <docSouls:name> xxx </docSouls:name>
      <docSouls:owner> xxx </docSouls:owner>
    ...
  <!-- ORIGINAL DOCUMENTS DOCUMENT CONTENT>
  <docSouls:originaldoc format="HTML">toto.html</docSouls:originaldoc>
  <docSouls:browsingdoc format="HTML">browse.html</docSouls:browsingdoc>

<!-- LIST OF PERSONALITIES AND DOCUMENT SERVICE REQUESTS>
<docSouls:endowmentdescription>
  <!-- LIST OF PERSONALITIES>
    <docSouls:personality id = " " name="" etc.>
      <docSouls:service id="" name="" etc.>
        <docSouls:provider name="" etc.>
          <docSouls:provider name="" etc.>
            <docSouls:service>
              ...
            <docSouls:personaligy>
              ...
    <!-- LIST OF SERVICE REQUESTS>
    <docSouls:service id="" name="" etc.>
      <docSouls:provider name="" etc.>
        <docSouls:provider name="" etc.>
      </docSouls:service>
    ...
  </docSouls:endowmentdescription>
```

FIG. 56

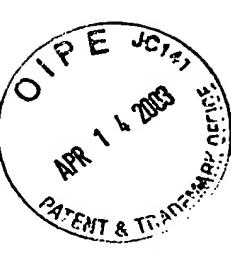
```
<!-- DESCRIPTION OF PERSONALITIES>
<docsouls:personalitydescription>
<docsouls:id> xx </docsouls:id>
<!-- all the fields AND service categories>
<docsouls:servicecategory id="" name = "">
</docsouls: personalitydescription>

<!-- NEXT PERSONALITY DESCRIPTION>
...
<!-- DESCRIPTION OF DOCUMENT SERVICE REQUESTS>
<docsouls:servicedescription>
<docsouls:id> xx </docsouls:id>
<!-- list all fields (e.g., "refresh period", category, etc.)>
<!-- list all providers>
<docsouls:provider id="" name = "">
<!-- list all strategies>
<!-- list status of services (original, frozen, mapped, etc.)>
<docsouls:servicestatus status="mapped" originalid="" originalserver="">
<!-- list ServerId if the service was not available on current server>
<docsouls:serverid>id</docsouls:serverid>
</docsouls:servicedescription>

<!-- NEXT SERVICE REQUESTS DESCRIPTION>
...

```

FIG. 57



```
<!-- DESCRIPTION OF PROVIDERS >
<docsouls:providerdescription>
  <docsouls:id> xx </docsouls:id>
  <!-- all the fields e.g. key, URL, etc. >
</docsouls: providerdescription >

<!-- DESCRIPTION OF DICTIONARIES >
<docsouls:dictionarydescription>
  <docsouls:id> xx </docsouls:id>
  <!-- all the fields, e.g. category >
</docsouls:dictionarydescription>

<!-- DESCRIPTION OF STRATEGY >
<docsouls:strategydescription>
  <docsouls:id> xx </docsouls:id>
  <!-- all the fields >
  <!-- all the dictionaries >
</docsouls:strategydescription>

<!-- RESULTS OF GLOBAL SERVICES >
<docsouls:globalservice id="Summary">
  <docsouls:link> xxx </docsouls:link>
  <docsouls:content> xxx </docsouls:content>
</docsouls:globalservice>

<!-- OTHER RESULTS (E.G. KEYWORDS, CATEGORIES, ETC. >
```

FIG. 58A



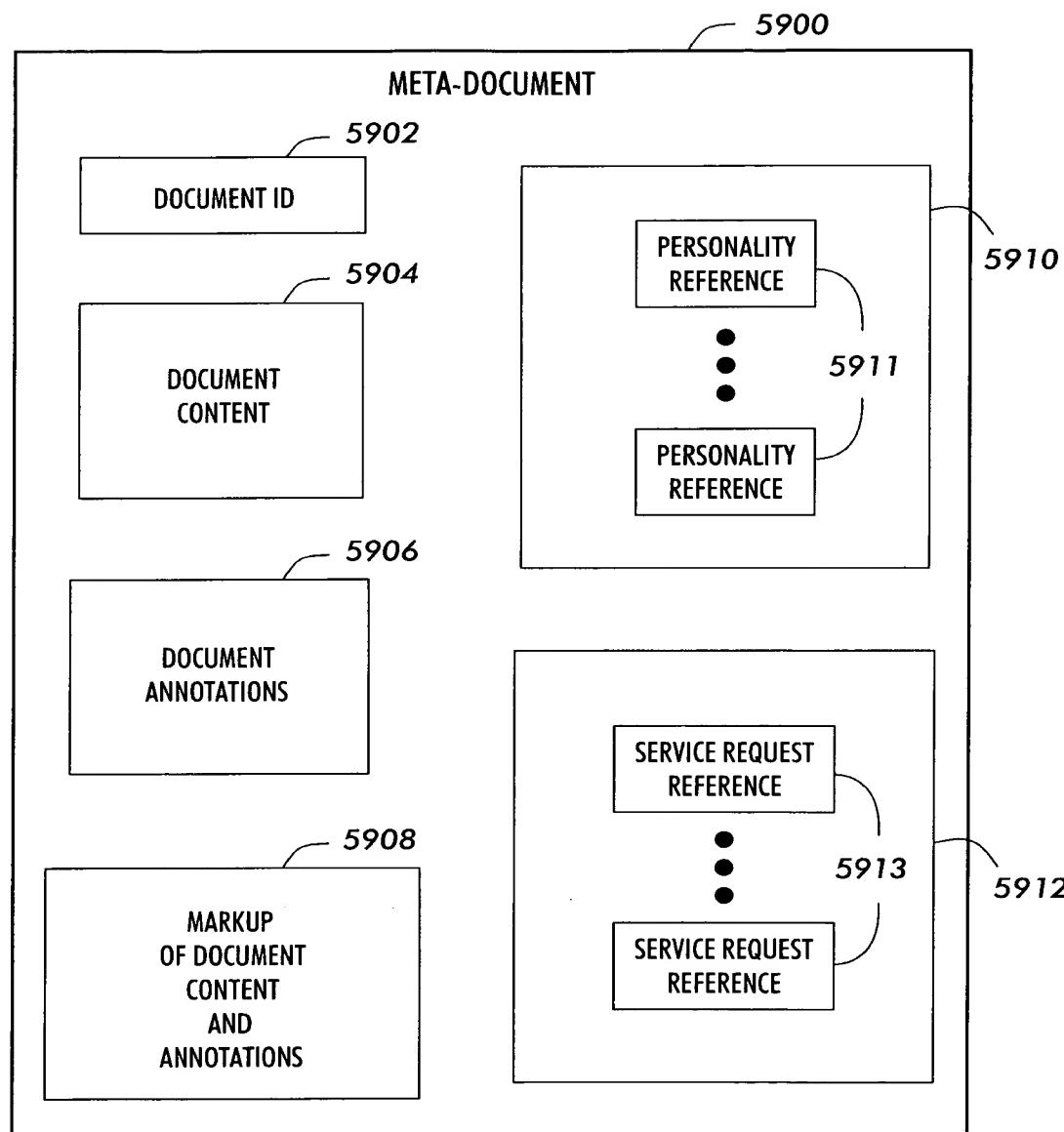
2176  
09/683,240

59/69

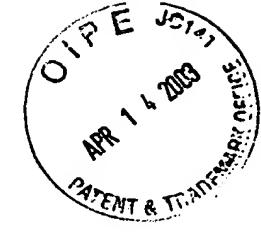
```
<!-- (OPTIONAL) DESCRIPTION OF ENRICHMENT PERFORMED AT OTHER DOCUMENT ENRICHMENT  
SYSTEMS>  
<docsouls:serverid>olderServerId</docsouls:serverid>  
<docsouls:personalitydescription>...</docsouls:personalitydescription>  
<docsouls:servicedescription>...</docsouls:servicedescription>  
  
<!-- (OPTIONAL) LIST OF RESULTS OF THE LAST ENRICHMENT >  
<!-- IDENTIFY DOCUMENT SECTION IN CASE OF MULTI-PART DOC >  
<docsouls:docpart>main_frame.html</docsouls:docpart>  
<!-- OFFSET WHEN ENRICHMENT APPLIES TO A SECTION >  
<docsouls:offset> 0 200 </docsouls:offset>  
<!-- STATUS OF SERVICE APPLIED ("ACTIVE" OR "FROZEN") >  
<docsouls:service id="StockQuote" status="active">  
<!-- ENTITY >  
<docsouls:entity>Xerox  
<!-- NORMALIZED FORM >  
<docsouls:normalize>Xerox</docsouls:normalize>  
<!-- LIST OF OFFSETS WHERE ENTITY IS FOUND IN TEXT RELATIVE TO SECTION >  
<docsouls:offsets start=35 end=40 />  
<docsouls:offsets start=232 end=237 />  
<!-- DICTIONARIES >  
<docsouls:dict>xxx </docsouls:dict>  
<!-- STRATEGIES >  
<docsouls:strat>xxx </docsouls:strat>  
<!-- LINK INFORMATION >  
<docsouls:link>xxx </docsouls:link>  
<!-- CONTENT INFORMATION >  
<docsouls:content>xxx </docsouls:content>  
</docsouls:entity>  
<!-- NEXT ENTITY >  
...  
</docsouls:service>
```

FIG. 58B

60/69

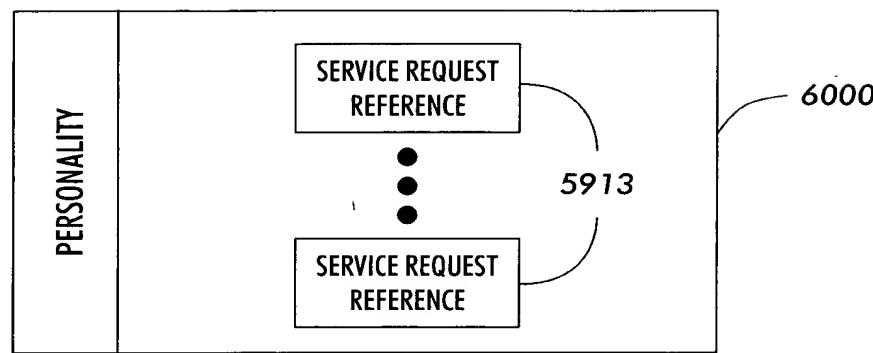


**FIG. 59**

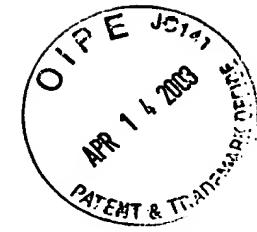


2176  
09/683,240

61/69



**FIG. 60**



2176  
09/683,240

62/69

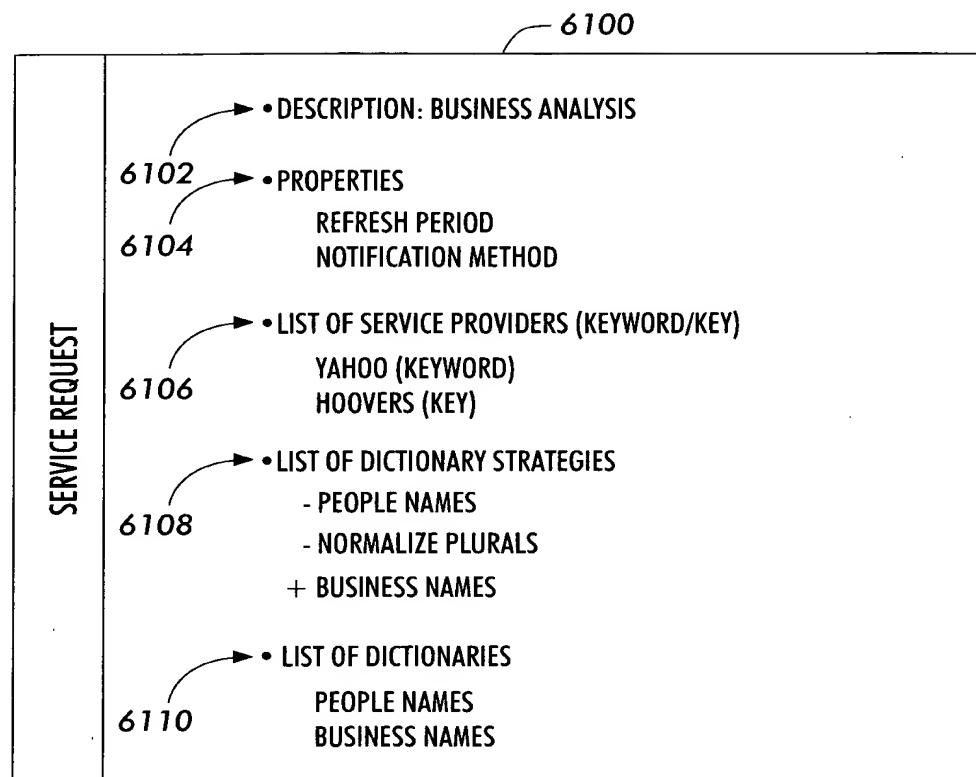


FIG. 61

63/69

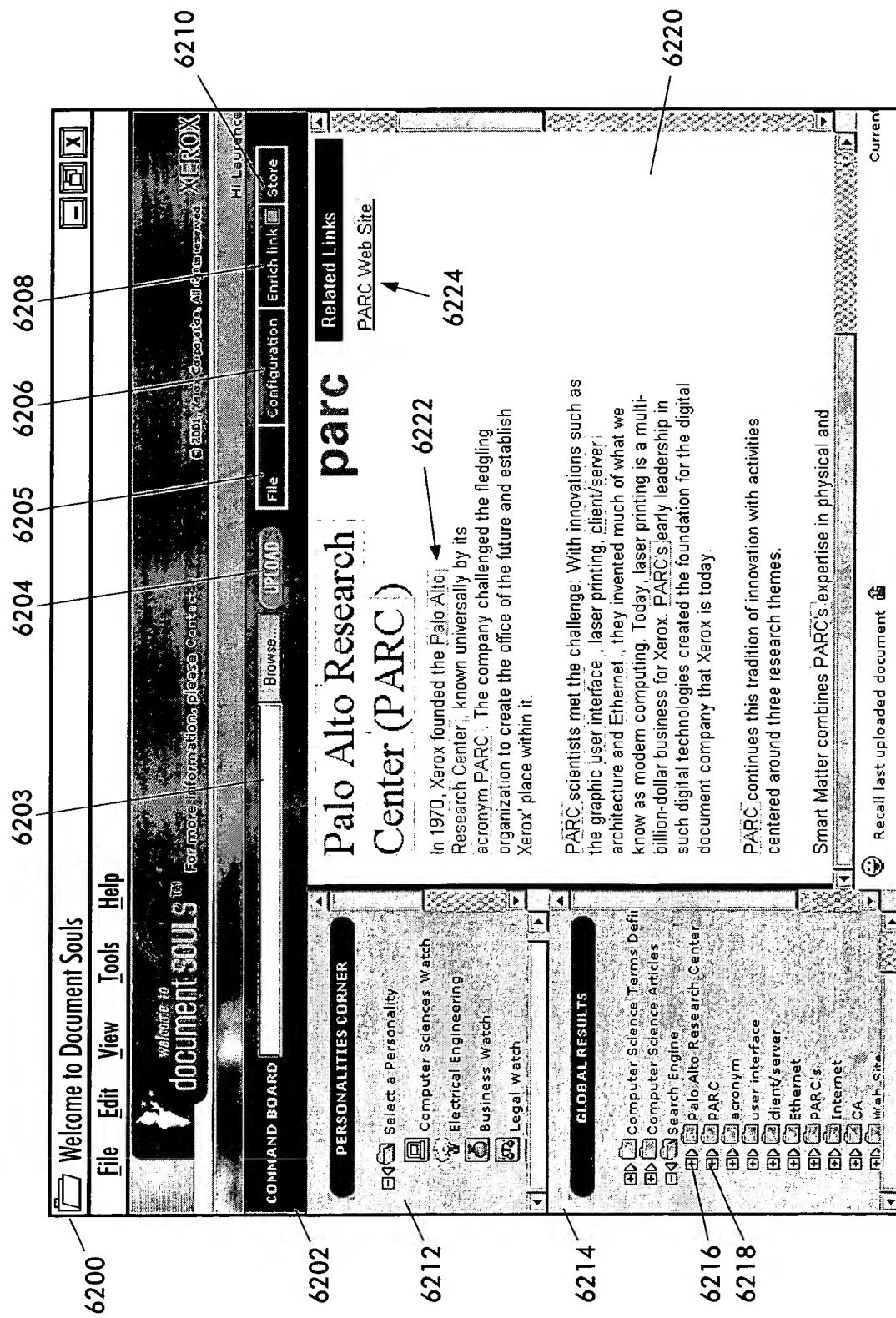




FIG. 63

65/69

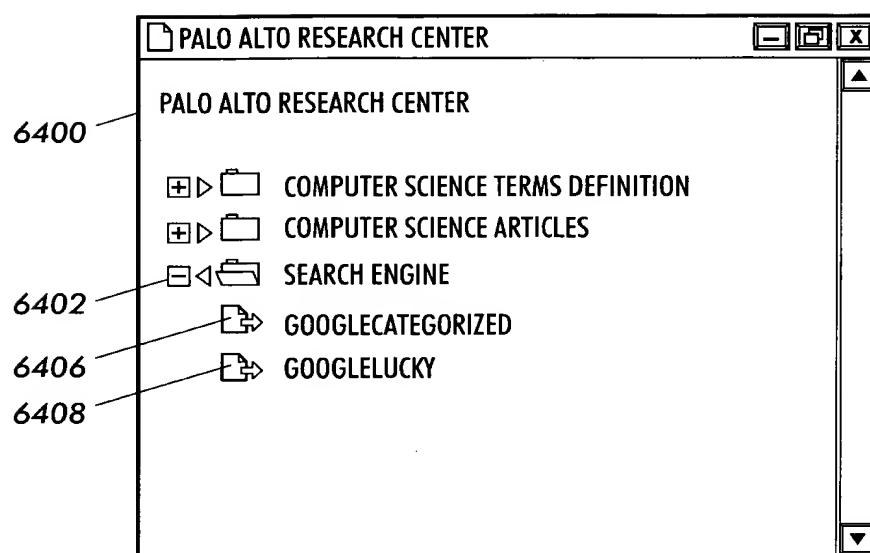


FIG. 64

66/69

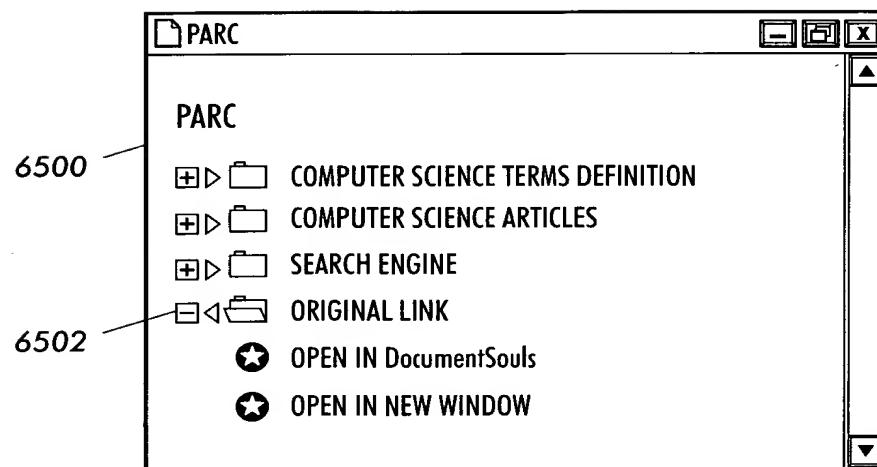
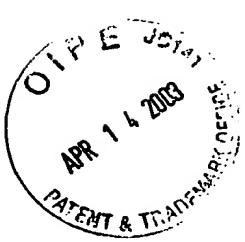


FIG. 65



2176  
09/683,240

67/69

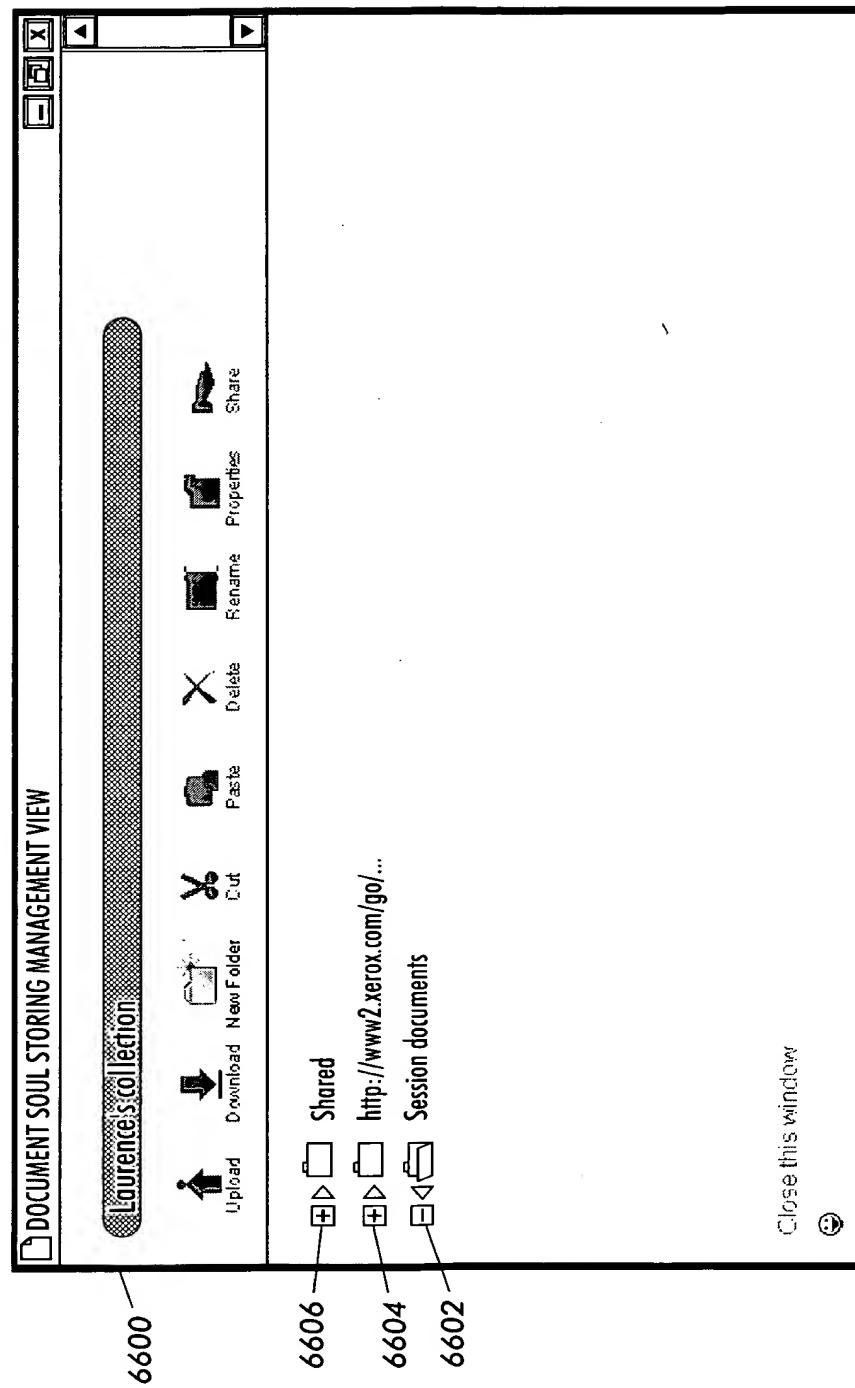


FIG. 66

68/69

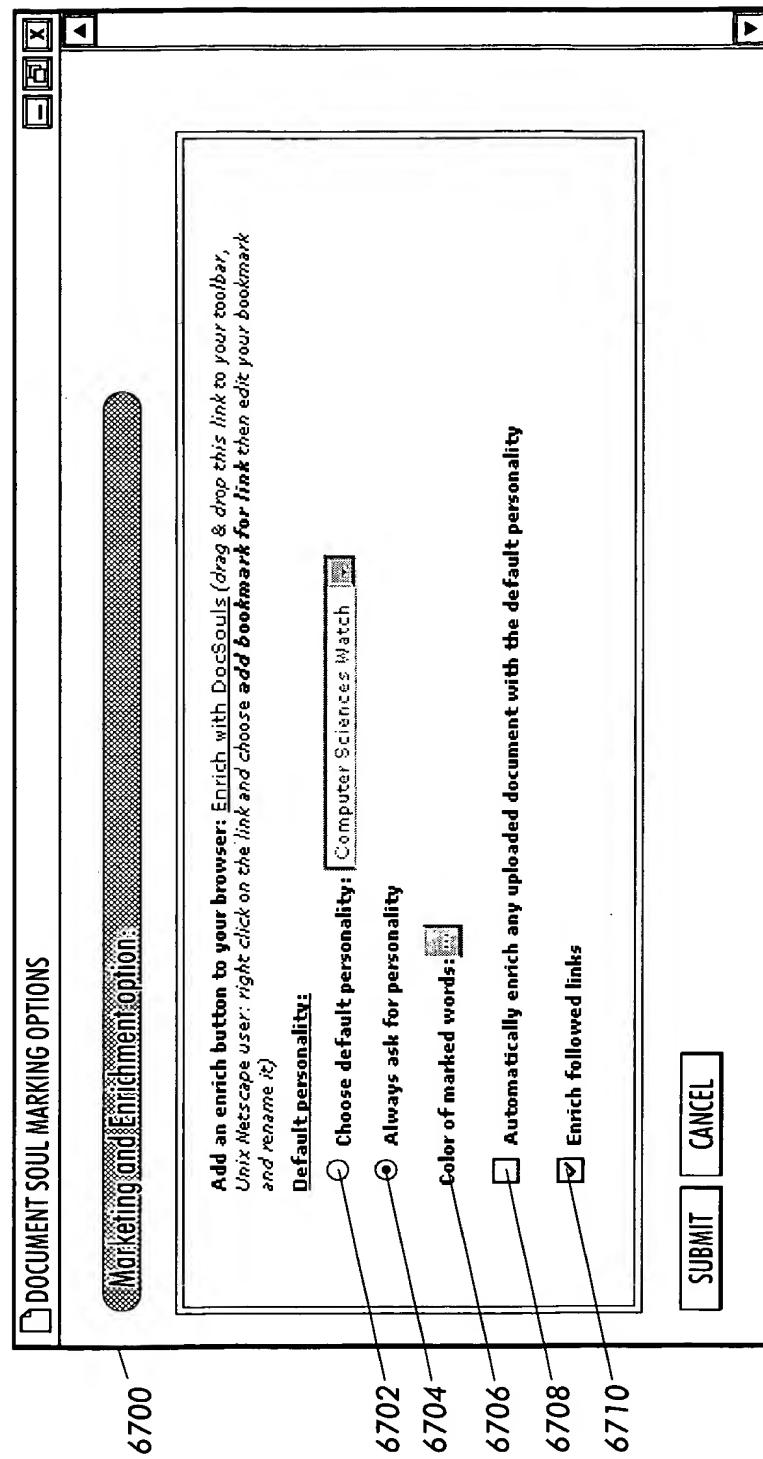


FIG. 67



2176  
09/683,240

69/69

6800

DOCUMENT SOUL SERVICES CONFIGURATION

List of available service categories

Abstracts

Name	Type	Status	Priority	Provider	Dict(s)	Refresh	Query
Abstracts from Compendex	STANDARD	Unknown	LOW	usPatent_office	list	none	LINK
Abstracts from Website Ab	STANDARD	Unknown	LOW	website_abstracts	list	none	LINK

6802

Industry Standards

Patents and IPs

Engineering Handbooks

Technical News

Technical Information

List of all Abstracts from Compendex additionalities

6804

Name	Category	Select
electricalEngineering	/Science/Technology/	<input checked="" type="checkbox"/>
names	?Society/People/	<input checked="" type="checkbox"/>

FIG. 68